RURAL-URBAN CONTEXTUAL DATA TRIANGULATION FOR INTERNATIONAL ENGINEERING PROJECT WORK

BY

ALEXANDRA TIMMONS

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Agricultural and Biological Engineering
in the Graduate College of the
University of Illinois Urbana-Champaign, 2021

Urbana, Illinois

Master's Committee:

Research Scientist Ann-Perry Witmer, Chair
Associate Professor Luis F. Rodríguez
Associate Professor Christy Lleras

ABSTRACT

Literature in the field of international engineering indicates that project failure is commonly a direct result of the exclusion of the community during important project and design decisions. This failure appears to be especially prominent for projects attempted in rural communities. The current emerging solution to address this community exclusion and resulting project failure is the application of contextual engineering. This involves a thorough and purposeful on-site study of community conditions with particular attention to social and physical conditions. However, this process is quite cost and time intensive for foreign engineers to engage with and in times of global travel restrictions it becomes impossible. This research acknowledges that the consideration of contextual factors is crucial to project success and studies the potential for a community's contextual conditions to be observed from abroad as a way to keep projects operational in situations where travel is not feasible. While rural community-specific data is largely uncommon, urban-level data is examined and an interpolation is performed. By selecting the three urban cities, where data is plentiful, in closest proximity by travel distance to the rural community in question and weighting them accordingly, it is hypothesized that a contextual engineering data triangulation analysis can be performed to approximate the conditions of the rural community. This process is tested through a case study of rural Tikonko, Sierra Leone and the results are compared against known community condition data gathered on-site. The results highlight the yet irreplaceable need for on-site interactions while still acknowledging potential avenues of usefulness for this type of remote data collection, as well as future work pathways.

ACKNOWLEDGEMENTS

Portions of the abstract and introduction were published by the author in 2020 in the American Society of Agricultural and Biological Engineers Technical Library under the title "Rural-Urban Contextual Data Triangulation Proposal for International Engineering Project Work".

An immense thank you to my advisors, Dr. Ann-Perry Witmer and Dr. Luis F. Rodríguez, as well as my committee member, Dr. Christy Lleras, for their constant support throughout this research journey. Their guidance was invaluable and will stay with me far beyond the constraints of this paper.

Thank you to the Contextual Engineering Research Group and BioMASS Lab for their immeasurable inspiration.

And an endless thank you to my family. To my aunt Bettina and my Yaya, thank you for cheering me on through every milestone of my life. To my brothers, Evan, Stephen, and Niko, thank you for your confidence in me. I hope to always be your strongest support system as you are mine. And to my mom, thank you for being the greatest mom I could ever ask for. I would not be half the person I am today without your unwavering love and encouragement. Sas agapó.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

The field of international engineering is a relatively new addition to the world of science and humanitarian aid. The process encourages the union of well-intentioned engineers, typically from Western countries, with communities in need of an engineering infrastructure solution. However, the resulting projects are not always deemed successful by all involved stakeholders, most notably the community members themselves. The literature indicates this is commonly a direct result of the exclusion of the community during important project and design decisions (Lee et al., 2018; Witmer, 2018b). This research acknowledges that understanding contextual conditions of a community, as well as including the voice of the community, are crucial to project success and hypothesizes that, in times where travel is impossible, this process can be attempted remotely. As not all members of the project design team of these international engineering projects formally consider themselves engineers by trade, this paper will refer to them as technical designers moving forward.

It is important to note that this research seeks to examine the potential for contextual engineering to be incorporated with big data through the lens of the hypothesized average user. This is a technical designer who understands the value of the contextual engineering, but does not grasp the foundational theories of the discipline. Thus, this technical designer would attempt a fully-remote process such as the one outlined in this research without truly understanding or being able to evaluate the potential limitations.

Even when travel is not possible, as is the current situation globally as a result of the COVID-19 pandemic, project progress is encouraged to continue. The methodology outlined in this research operates with the understanding that speed and simplicity are favored over a complex sociological analysis process, as these are the conditions likely to be preferred by the average technical designer (Litchfield & Javernick-Will, 2014). This is inherently opposed to a core tenet of contextual engineering which states on-site face-to-face interaction with project clients is the key to understanding community conditions and building the foundation for the best chance of project success. However, technical designers are currently forced to operate remotely, and will likely continue to face similar hurdles in the future. Therefore, this research is a step towards expanding the efforts of the discipline, testing its application and limits, as well as attempting to address this current gap.

By studying potential incorporation of remotely available data, this research creates and examines the efficacy of a contextually-informing remote-data process to be utilized in international engineering work. To combat the lack of rural community-scale specific data, this research examines the potential use of urban-scale data triangulation to inform rural project design. Also significant to this proposed project process is the incorporation of comprehensive community feedback throughout the project lifetime. The differences between the generally accepted international engineering project design process, and the proposed modified international engineering design process are visualized (Figure 1.1). Studied in this work is the potential to partially replicate the on-site assessment with a combination of remotely accessed contextual data supplemented with conversations with the community through the life of the project. Additionally, the suggested process emphasizes the cyclical and iterative nature of successful project design, as well as the potential for the process to end before the formal project conclusion if a satisfactory agreement cannot be reached.

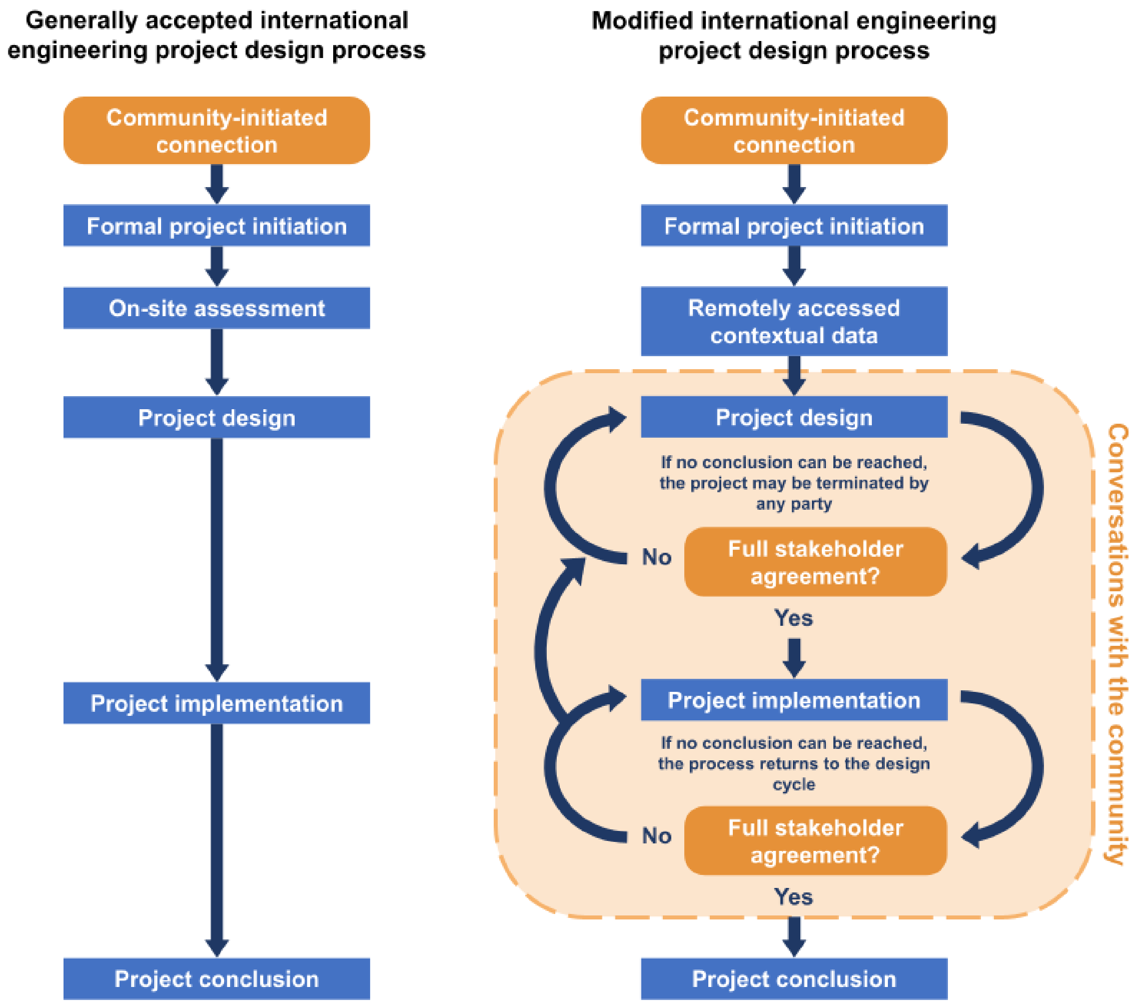

Figure 1.1: Comparison of the generally accepted international engineering project design process to the proposed modified international engineering design process. Rounded, orange boxes delineate a formal involvement of community member input in contrast to blue boxes representing internal operations on the end of the engineers. The dashed orange box surrounding the project design and project implementation elements of the modified process is representative of conversations with the community occurring regularly throughout the process. Notably the on-site assessment box of the generally accepted process is replaced with remotely accessed contextual data supplemented with conversations with the community in the modified process. Additionally, the process format is adjusted from linear to cyclical and iterative.

The research is guided by the following questions:

1. **How can we better identify relevant rural community conditions and contexts from a distance?**

   This question is particularly relevant for technical designers who are limited by time, cost, or travel restrictions; such as the case generated by the COVID-19 pandemic (CDC, 2020; Chinazzi et al., 2020).

2. **Assuming that urban communities are generally more researched than rural communities and therefore have larger and more up-to-date data pools, how can existing urban data most productively inform rural engineering infrastructure design technical designers?**

   While there does exist a smattering of data sources focused on select rural communities, these are typically case studies conducted for individual research agendas. As such, there is not much room for generalized conclusions or global applicability. Thus, urban-level data is selected for its broader global availability. Data acquisition is elaborated on in Section 4.1.

3. **Which available metrics or indicators provide the most telling information to help inform rural international engineering infrastructure design?**

   This search is based on the necessary data recommendations put forth by Witmer, 2018a following the established work of Contextual Engineering.

4. **What factors differentiate rural communities from urban**

communities and how do these factors play a role in engineering infrastructure design?

While the methodology will be interpolating rural conditions from urban data, there are factors that differentiate rural and urban communities that must be considered.

CHAPTER 2

LITERATURE REVIEW

The prevalence of international engineering projects has grown immensely since the initial conception of some of the largest players in this field in the late 1990s and early 2000s. These international engineering projects have gone by many names, including humanitarian engineering, sustainable community development, and service-learning engineering among others (Duffy et al., 2000). At the center, all refer to the involvement of foreign engineers assessing, designing, and implementing an engineering infrastructure project in response to a community need. While the field continues to develop, its relative novelty leaves room for continued growth and adapted practices in order to ensure the mission statements of these projects continue to be met.

One of the clear challenges of international project work is adapting to the cultural and contextual differences that present themselves at various steps throughout the project process. Even obstacles as predictable as language barriers can impede progress towards a successful project (Calefato et al., 2010). But the extra steps necessary to gather data on social and physical conditions and then analyze and incorporate these considerations into the project process can make the barrier to entry seem intimidating for those unfamiliar with the process and/or the benefits. Additionally, gathering this data on context-informing conditions benefit immensely from on-site investigations (Ridgway, 2015), which are not always possible to conduct, such as the case with the global travel restrictions necessitated by the COVID-19.

It is also observed that there is a growing global affinity towards the use and reliance on big data (Lohr, 2012; Madden, 2012). As there is a need for the principles of contextual engineering to become more accessible and as there are a plethora of databanks containing an abundance of local data, this research attempts to integrate the two and evaluate both the effectiveness of this process and any potential obstacles along the way. In addition, this research proposes that rural engineering infrastructure design projects conducted by technical designers not native to the country of the project's clients can benefit from a modified project process by incorporating client feedback throughout the life of the project. While this consideration won't be directly examined in this particular study as it fell beyond the scope, the implications will be considered throughout this research and the suggestion is a consideration for future work.

2.1 Examination of Chosen Research Areas

To explore the hypothesis that context can be learned remotely and attempt to answer the guiding research questions, an examination of the existing literature has been broken down into core categories; global significance of rural communities, data utilization methods currently applied in rural international engineering design, and relevance of contextual data and information beyond technical data. In depth examinations of each of these areas will aid in exploring the deficiencies of existing methodologies applied in international engineering projects and the gaps this research aims to fill.

2.1.1 Global Significance of Rural Communities

As previously mentioned, the field of international engineering projects has grown rather quickly in the roughly two decades since it formally came to be. While many organizations have emerged in this time, the common theme between them all is clear, the projects are intended for disadvantaged communities in need, which predominantly applies to rural communities. With statistics like "68% of the world population projected to live in urban areas by 2050" (United Nations, 2018a), it may seem that rapid rural out-migration is inevitable. Yet, in truth, the global rural population is still on the rise and is predicted to rise for a few more years until the decline begins (United Nations, 2018a). And even with a decline, what many sources highlight when referencing this out-migration is the generalized rural desire for "opportunity" while failing to acknowledge the many factors that impact the viability of this opportunity as well as the context in which opportunity is desired (FAO, 2018). While these reports address the benefits of rural out-migration and acknowledge that ultimately, "government policies and programmes in areas of both origin and destination play a key role in determining the final impact of migration" (FAO, 2018), the argument in favor of rural out-migration boils down to a predominantly economic one. The generalization of a net-positive economic benefit has the potential to be damaging to rural communities who do not share the same values as those being examined. For one, it fails to truly address the challenges beyond economic disadvantage, the threat of a loss of social capital, and sense of place, is a relevant byproduct of rural out-migration (Wang et al., 2020), as well as ignoring the portion of rural populations that have no interest in migrating at all, regardless of potential economic benefits. Additionally, this pattern of thinking tends to neglect a number of motivating migration factors beyond perceived monetary gain that can be achieved within rural communities, if supplied the right resources (Balodi et al., 2015). Ultimately, the potential outcomes of investing in rural communities, as opposed to encouraging their out-migration, vary significantly according to context (Deotti & Estruch, 2016), but it stands that investing time and research into the sustainability of rural areas is a worthwhile objective that is and continues to be a relevant and worthwhile objective.

A challenge faced by researchers studying rural communities is that there is no singular global definition for the term. Rather, "rural" and "urban" are defined uniquely by each country and a variety of metrics are utilized across the board to do so. These metrics include population, population density, and infrastructure development levels, which takes characteristics such as the presence of paved streets and electric lighting into account (Ritchie, 2018). Even these designations can be misleading in their own ways. Bo City in Sierra Leone, for example, the second largest city in Sierra Leone with over 170,000 inhabitants, does not have working electric lighting throughout the city (Statistics Sierra Leone, 2016). This anecdote emphasizes the importance of utilizing regional definitions as opposed to attempting to make a universal definition, but also highlights the challenge these nuances and specificities create when making comparisons on a global scale, as this research will examine. Table 2.1 displays information sourced by the United Nations' World Urbanization Prospects demonstrating some of the immense variety between nations' definitions of "urban" (United Nations, 2018b).

Table 2.1: Sample of author selected nations and their national definitions of "urban" as determined by national statistical offices, the latest available census, and UN estimates. These nations were selected to sample a variety of definition types as well as provide some insight into countries that will be further discussed in this paper (United Nations, 2018b).

Country	National definition of "urban" as determined by national statistical offices, the latest available census, and UN estimates
Anguilla	In the absence of more detailed information the entire population is considered urban.
Bangladesh	Localities having a municipality (pourashava), town (shahar) committee or cantonment board. In general, urban areas are a concentration of 5,000 inhabitants or more in a continuous collection of houses where the community sense is well developed and the community maintains public utilities, such as roads, street lighting, water supply, sanitary arrangements, etc. These places are generally centres of trade and commerce where the labour force is mostly non-agricultural and literacy levels are high. An area that has urban characteristics but has fewer than 5,000 inhabitants may, in special cases, be considered urban.
Greece	Municipalities and communes in the largest population centre with 10,000 inhabitants or more, plus 18 urban agglomerations as defined in the 1991 census: Greater Athens (Athínai), Thessaloniki, Pátrai, Iraklion, Vólos, Chania, Irannina, Chalkida, Agrino, Kalamata, Katerini, Kerkyra, Salamina, Chios, Egio, Rethymno, Ermoúpolis and Spárti.
Iceland	Localities with 200 inhabitants or more.
Sierra Leone	Towns with 2,000 inhabitants or more.
United States of America	Densely settled territory that meets minimum population density requirements and with 2,500 inhabitants or more. A change in the definition for the 2000 census from place-based to density-based affects the comparability of estimates before and after this date.

Of course it is understandable why such categorizations and assumptions are made, in order to compare any of these valuable global data, certain standardizations must be adopted (United Nations, 2017). But it is also evident that such broad definitions have the potential to let many communities of people slip through the cracks of analyses and therefore decision-making considerations, further stressing the need for contextual considerations in fields of work across the globe.

2.1.2 Data Collection and Utilization Methods Currently Applied in Rural International Engineering Design

Regarding current processes in use to collect data for international engineering projects, a brief examination of some of the largest organizations in this field will be conducted. Engineers Without Borders, likely the most transparent organization, utilizes a step-by-step series of documents to guide their volunteers through the project process. The on-site work is broken into three phases, each requiring their own trip, assessment, implementation, and monitoring. After the chapter's project application, the next documented step outlines data collection to be conducted on-site. There is no documentation to indicate that conversation between the community and the engineers has occurred since the project conception, leaving the project assessment trip to be the sole opportunity to gather data and information about the community. While there is language that indicates considerations should be given to non-technical relevant data, such as community relations and community priorities, for example, (EWB, n.d.) it is very limiting in its expectations of data collection and does not outline what is to be done with this information, let alone in case of initial conflict, such as community priorities not aligning with EWB expectations. However, the organized requirement of an initial assessment trip to be completed before project construction begins does demonstrate consideration of potential unanticipated conflict or obstacles.

From what can be publicly sourced, it seems Engineers Without Borders does offer a handful of other internal resources, but these resources are not easily accessible, nor are they widely promoted. However, their existence is evident of an acknowledged interest in contextually-minded thought processes and data utilization. For example, Engineers Without Borders requests their project teams submit a "Lessons Learned" project summary after the conclusion of the project. These short-form reports aim to collect and categorize brief feedback statements about elements of projects that students wish they had done differently, or known of ahead of time. The type of feedback that EWB recommends students provide include "communication tips" and "culturally significant comments", in addition to more technical considerations such as "ease of access to in-country materials and equipment" and "technical comments learned about the region's current infrastructure" (EWB, 2019). Since it's conception in 2016, this database has amassed over 3000 "lessons learned".

Regarding data collection specifically, there appear to be a number of comments that highlight the need for pre-travel contextual considerations. For example, students working on a project in West Africa noted, "Many water sources go by many different names… [in the future we must] record all of these different names" and "get a detailed understanding of the methods that the contractor will use before agreeing to hire... we had to bring in a second contractor... because the first used a novel method with a technique that is not as widely accepted in the field". A group working in East Africa noted, "We did not consider flooding in our designs and failed to ask the community any questions about the rains that did not [pertain] to drinking water". Students working in the Caribbean documented that while they wanted to collect data on the entirety of the communities water supply, "the water lines and service to houses are regularly added and closed with little to no supervision or documentation... we should have worked closely and very clearly with the community to make sure the community took ownership and contributed to, instead of working against, the mapping process" (EWB, 2020a). What these lessons share in common, is their lack of communication with their client communities. As a result, their collected data was mislabeled or insufficient, and their design plans were incompatible with community need. This database has great potential to be a resource for future teams. For example, if readers could respond to lessons and bump beneficial lessons to the top of the page, or if studying the Lessons Learned database was a required element of the project process, there may be more potential for this knowledge to be applied.

Similarly, in 2013, EWB spearheaded an initiative known as the Planning, Monitoring, Evaluation and Learning (PMEL) Program with the intent of directing resources towards evaluating if EWB projects were truly functional as intended. Hand-in-hand with this program is a series of Impact Reviews that are intended to allow EWB "to determine if our work had any lasting impact in the community" (Martindale, 2013). While the reports have the benefit of a more thorough analysis than the previously discussed Lessons Learned, they are much more susceptible to funding and manpower concerns which is evident in the program's current state. Only one country is reviewed per report, and while these were intended to occur annually, it appears only four total have been conducted, with the most recent Impact Report published in 2017 (EWB, 2020b). Additionally, while the reports made note of interviews and conversations had with the recipient communities themselves, no first-hand community input was included in these analyses of impact. In a 2015 external review, the Impact Analysis Director of the PMEL Program at the time acknowledged the promise shown since the program's creation, but due to the "large administrative and reporting burden, that the resulting products are not as usable and accessible to project teams as desirable, and that it is difficult to determine if learning is actually being put back into their processes" (Ridgway, 2015).

Other prominent organizations shared many of these same traits. Engineers Against Poverty, for example, advocates for the elimination of poverty through sustainable infrastructure and improved social equity for impoverished communities, yet state they intend to achieve this goal by drawing on their own research and by "producing innovative knowledge products", with no acknowledgement of the community members themselves (Engineers Against Poverty, 2021). The intention is not to say that there are no successful projects among these organizations or that these projects are not well-intended, but that ineffective strategies are not uncommon in the field, and that community-focused solutions are not always the genuine outcome of project work.

2.2 Understanding Contextual Data: The Five Critical Influences

Contextual data is a term representing the non-technical data that plays an important but often neglected role in understanding a community, coined by Dr. Ann-Perry Witmer (Witmer, 2018b). In order to create sustainable and successful engineering solutions, the needs of the community must not only be met, but be the primary driver (Jordan, 2015). While the definition of a project's success is dependent on which stakeholder is consulted (Schreiber et al., 2019), it stands that unless the subject client community is satisfied with the outcome of the project as the primary users, a project should not be deemed successful. Therefore it is crucial for engineers to make every reasonable effort in understanding and supporting the stated needs of the client community through their project design (Walters & Javernick-Will, 2015). While there are a variety of ways to shape these non-technical influences into manageable and efficient categories for analysis, this research will proceed using the five critical influences outlined by contextual engineering: cultural, economic, educational, mechanical (formerly technical), and political (Witmer, 2018b).

These influences intend to address the vast array of factors that contribute to a community's unique identity, divided into manageable categories. Each of these influences has a practical effect on project design decisions (Table 2.2), but not all are always present in each project. It is necessary that the technical designer evaluates each community individually to determine the strength and relevancy of each influence in order to understand which are dominant. This evaluation is most directly accomplished through the use of Witmer's Predictive Tool (Witmer, 2018b), which was created for this purpose and can be found in Appendix B.2. However, this tool requires on-the-ground evaluation, which, as established, may not always be possible. Through a thorough understanding of contextual engineering, this research seeks to understand these same critical influences and examines the potential for them to be addressed from abroad.

Table 2.2: Contextual data is categorized into five critical influences; cultural, economic, educational, mechanical, and political. These five influences play a crucial part in understanding a community and designing international engineering projects effectively

	Working Definition	**Practical Example**
Cultural	This influence acknowledges data and considerations regarding a community's belief system, spiritual values and other cultural considerations	A community may not culturally agree with a particular building material or practice
Economic	This represents the economic need or lack thereof of a community	Economic restrictions may dictate the scope of the project design, or potential labor hires
Educational	The educational influence measures the degree to which a community values the acquisition and dissemination of knowledge	Community with high education completion rates may imply a stronger community interest in being a part of the design process
Mechanical	A community's mechanical aptitude is determined by their self-sufficiency	A strong community interest in self-sufficiency may allow for greater project complexity and more supported maintenance potential
Political	The political climate of a community may be dictated by factors such as which members are allowed to participate in politics, and hierarchical living arrangements, among others	An egalitarian society may not have a preference on project location, while a politically divided one may have a strong preference

2.2.1 Importance of Engineering Context

Understanding the social and physical conditions that contribute to a community's identity are crucial to designing a project appropriate for their unique needs and circumstances (Witmer, 2018a). However, even when these contextual elements are evident, they are not always incorporated into project design decisions. For example, a student engineer struggled to convince their mentor that a typical US barrier wall design would not be appropriate for their specific project due to an obvious lack of appropriate construction equipment within the community (Lee et al., 2018).

However, not all contextual conditions are as immediately apparent. Lee et al., 2018 also observed that students found themselves surprised by how commonly and frequently international engineering project participants "treat 'communities' as undifferentiated entities that share unified goals and needs" While these contextual differences were observed, the students did not have the education to properly acknowledge them within their design. In one example contained within this study, this led to a "severe rift among community members" as non-community members were hired to work on the construction elements of the project, as opposed to residents as previously promised. After an anthropologist was brought on one of the trips to the community with the team, the project impact was reevaluated and shortly after, closed. A student engaged with the project stated the following:

> As much as we want to say it's for the communities, it's really about the students for the club. Because if you think about it realistically, okay

> we're doing a couple assessment trips, but you don't really gauge full community needs through a few weeks. I'm a firm believer in that. So, okay, you can try, maybe do a few small projects, and if you're ambitious maybe you'd get a large project, but as far as long term success, I wouldn't really say it's as much for the community as it is for the student. (Lee et al., 2018)

While this story is anecdotal, it does reveal a sentiment that is shared among others in this line of work. Determining a balance between student and community engagement is tricky. Too little focus on ensuring positive volunteer experiences could dwindle necessary numbers of student involvement required for the organizations to function. Too much emphasis and projects end up like the example shared above, with not enough consideration allocated for the true needs of the community. This work aims to operate in that balance by acknowledging that full community needs cannot be determined without proper education on the importance of contextual data collection.

During the author's research travel to Sierra Leone in the summer of 2019, similar stories of failed engineering infrastructure designs were recounted by community members. In one such case, the author was interviewing community members about their local water supply. The community leader angrily shared that they received foreign aid that resulted in a large gravity-fed water system constructed in the middle of their village (Figure 2.1). The imposing structure had not been functional for some time as no one in the community had the knowledge on how to repair it and they were not successful in their attempts to get a hold of the original engineers. Additionally, the community desired the hand pump most neighboring communities had as it was familiar to operate and within their abilities to repair should something go wrong. They expressed that receiving the gravity fed system instead made them feel ignored and discouraged them from being willing to work with foreign aid groups in the future.

Figure 2.1: Photographs taken by the author during her travel to Sierra Leone. The hand pump water system displayed on the left was seen commonly in villages across the Tikonko Chiefdom and was easily operated by community members of all ages. The gravity fed water system on the right was located in one particular village in the Chiefdom. This village expressed disdain for the system as it no longer functioned and wasn't the familiar hand pump system they desired.

In another example, while driving through Bo City, our local travel partner pointed out the street lights we passed as we drove through the heart of the city (Figure 2.2. The street lights were installed as part of a government-funded project to make the streets of Bo safer at night. The lights were solar charged and powered by two battery boxes just below the light itself. While the infrastructure functioned as designed initially, many of the street lights are no longer operational. It turned out that many people saw a more pressing need for the batteries, and scaled the poles to remove them and use them to charge their cell phones instead.

Figure 2.2: Photograph taken by the author during her travel to Sierra Leone. The image depicts one of many solar powered street lights in the city of Bo. In the cropped version of the photo on the right, the two battery boxes storing power for the street light are more clearly visible.

Neither the gravity fed water system nor the solar powered street light were unsuccessful as a result of poor technical design. Rather, the contextual needs of the community were not well understood before the project was implemented, resulting in the failure of both systems. This research will address this observed deficiency by creating and testing a methodology to search for this necessary context within remotely-accessible data. The intent is that this exploration will aid rural international engineering project participants in incorporating contextual information into their project designs when travel to the community in question is not possible. As the review of the existing literature has shown thus far, this is a relatively new field and as such, there is not much published in this sphere yet, let alone methodologies that examine how to improve this work. However, it is evident that there exists room for improvement in the existing project process, especially with regards to the impact of contextual data on project design.

2.3 Conclusions

After an in-depth review and analysis of the existing literature, the need for improved rural international engineering methodology is still apparent. It is also evident that engineering work that employs contextual mindsets to every step of the project process, from conception through monitoring, is a highly beneficial addition, but lacks the breadth of research necessary to apply it in any variety of situations. With minimal work accomplished in this field thus far, this paper aims to both continue to validate the need for further related literature, as well as explore a methodology that may begin to address some of the existing shortcomings observed in previous projects.

Existing urban-centered data triangulated around the rural community in question has the potential to inform community context. By creating a guideline for contextual data interpretation and resources that direct technical designers to city-level data, this remote data has the potential to more accurately address their rural client's needs, as well as providing the much undervalued benefit of making the user familiar with contextual thought. While some scholars have indicated that the perception of shared values between rural and urban residents is split, leaning more dissimilar than not (Parker et al., 2018), this research will be examining urban-level data for its ability to accurately reflect local rural conditions as urban-level data is the smallest granularity available. The research will evaluate if this is a sufficient method and to what degree.

By addressing the importance of utilizing local data, the research hopes to additionally emphasize the significance of utilizing a tool like Witmer's Predictive Tool which puts the client's needs at the forefront of the project and helped guide the foundation of this study.

CHAPTER 3

METHODOLOGY

This chapter overviews the data sources utilized, the type of data collected from the source, and how that data was reviewed and analyzed for a case study of a rural community in Sierra Leone.

3.1 Quantifying Contextual Data

As previously defined in Section 2.2, contextual data is the information required to paint an illuminating picture of a community's social and physical conditions. This collected data is then categorized into the five critical influences through the use of Witmer's Predictive tool (Witmer, 2018b) providing technical designers with a guided lens through which the relevant conditions are more clearly understood. As one can imagine, contextual data is often qualitative. The data collected is sourced from observation and visual cues, filtered through the lens of personal experience, which encourages the observer to take stock of their surroundings in a purposeful way. Witmer's tool then takes that largely qualitative observational data, organizes it through the use of Likert scales, categorizes, weights, and compiles the measured indicators according to their ability to inform the various influences, and returns a community score. This score is a quantified representation of what influence or influences are most dominant in a particular community. This information then provides some contextual structure to aid technical designers, particularly during the design and implementation process.

The data in its raw form is not directly translatable into context-informing data. At this point, given the novelty of a contextually-minded approach, there does not exist a system or methodology to guide this process. Additionally, the nature of this research, with its intent to be applicable for any global rural community, also implies that any one-to-one process will not be possible. The variations in the raw data can be accounted for to an extent, but there are limitations. Thus, this process will attempt to predict the typical data one can expect to find available, and outline the process for which that data can be interpreted in a contextual manner. The direct application of this process will be made later in this paper through a case study analysis of rural Tikonko, Sierra Leone. The collected and quantifying contextual data gathered remotely will then be compared against known on-the-ground conditions from an earlier site visit. This case study analysis is expanded upon in Section 4.3.

For the purposes of this research, the remotely available data is evaluated only for its ability to predict contextual conditions, not technical conditions such as topography or intended user population.

3.2 Tikonko, Sierra Leone: A Case Study

This section details the process of conducting a contextual analysis of a rural community through the use of remotely-accessible urban-level data. The baseline known conditions are established, the data collection process is reviewed, general contextually-relevant indicators are established, the available data is parsed through in search of matching indicators, the data is weighted by geographic distance, and critical influence scores are computed.

3.2.1 Introduction to the Case Study

This case study considers the rural village of Tikonko, Sierra Leone as the focus of a hypothetical international engineering project. The village is the capital of the Tikonko Chiefdom, and located in the Bo District of Sierra Leone's Southern Province. The village itself is home to a few thousand of Sierra Leone's nearly 8 million citizens (World Bank, 2020). For the purposes of this case study, it is postulated that the technical designer will know or be able to gather the relevant technical conditions and data, but contextually will not know much about the rural community, as they have not visited the community or spoken with any community members other than perhaps one or two briefly during the project initiation phase in a virtual capacity.

The case study considers the methodology according to the country-specific conditions at hand to create a contextual understanding of rural Tikonko. The data is sourced from the Sierra Leone data portal of the Africa Information Highway (AIH) aggregated open source data network (Africa Information Highway, 2020a). Specifically, the selected data is sourced from the national statistical office of Sierra Leone, Statistics Sierra Leone, and the specific indicators utilized as the methodology progresses will be largely unique to this dataset. Further information about the selected data source can be found in Section 4.1.

The data collected and analyzed from this case study will be compared against the authors contextual evaluation which was completed on-site and evaluated for effectiveness in Chapter 4. It is important to reiterate that this methodology is not attempting to find the absolute most effective way to interpret remote data in a way that portrays an accurate contextual understanding. Rather, the methodology is attempting to be efficient, not too time intensive, and conducted under the assumption that the user of this methodology is relatively new to the concept of contextual engineering. Throughout this case study, these base conditions are kept in mind, and used to guide a more realistic use case of the process.

While Sierra Leone does not collect data specific to smaller rural villages such as Tikonko, information and data about nearby larger urban areas is documented in a variety of its national statistical data. To determine the urban cities to be used for this methodology, first a list of all urban cities with available data was compiled. From this list, travel distance from Tikonko to the available cities was calculated using OpenStreetMap and taking all transportation methods into consideration, including driving, biking, and walking. For Tikonko, circled in red, the three urban areas in closest proximity to the rural community are Bo, Kenema, and Pujehun, each boxed in blue (Figure 3.1). Tikonko, Bo, and Pujehun are all part of Sierra Leone's Southern Province, while Kenema is located in the Eastern Province.

Figure 3.1: The rural village circled in red is the capital of the Tikonko Chiefdom located in the Bo District of the Southern Province of Sierra Leone. While there is not readily available data specific to Tikonko for the purposes of remote access project design and community understanding, there is a wide selection of data available for nearby urban Bo, Kenema, and Pujehun, boxed in blue (Nations Online, 2014).

The hypothesis is made that the closer geographically a rural community is to another neighboring community, the more similar those two communities are likely to be. We recognize that a variety of factors other than travel distance may influence an urban area's potential degree of correlation with a rural area, such as shared commerce, cultural makeup, and political affiliations. In Sierra Leone, for example, each province is divided further into districts, and further again into chiefdoms. Each district has a representative in the national parliament, while each chiefdom has a local leader and council, all with a variety of political views (Reed & Robinson, 2013). However, as previously acknowledged, the introduction of the numerous additional sources that would be required for this much more extensive analysis would contradict the ease-of-use intentions of this methodology. The efficacy of this approach will be evaluated in Chapter 4 and the potential for these additional sources to be factored into the analysis will be revisited in Section 5.1.

The overall framework of the methodology is as follows. Once the rural village is known, the corresponding country database is combed for datasets containing urban-level data. External known contextually-informing indicators guide the selection of useful indicators from the available remotely-accessible data. The individual indicators are scored following the same scoring methodology as Witmer's Predictive Tool. After appropriately weighting the data by contextual importance, one score for each critical influence will be generated for each city. The city scores are weighted by travel distance from the rural village. The generated score is compared against the known data collected during an in-person site visit. Efficacy of the methodology is evaluated.

3.2.2 Gathering the Data

As our case study considers a rural village in an African country, the open source AIH database is sufficient. A free account was made with Open Data for Africa to facilitate this data collection process. Within the Data Catalog of the Sierra Leone Data Portal the available datasets were sorted in two ways to pull the desired urban-level information. Using the Dataset Browser, first the data was filtered by Regions as the data available at the regional level is likely to disaggregate further into the urban level. The online interface separated the data into four regions of Sierra Leone, Eastern, Northern, Southern, and Western Area. These regions are representative of the four Provinces of Sierra Leone of the same names. However once an individual data set is accessed it allows further disaggregation into about 15 cities in addition to the country level averages. The available datasets at this level were then further examined to confirm they contained urban-level data. Those that did were flagged for reference.

It is at this point in the methodology that the three closest urban cities to the rural community in question can be selected, as the cities with available urban-level data are now known. As discussed previously, the distance from rural Tikonko to each available urban city was calculated using OpenStreetMap and considering all transportation methods and the three cities geographically closest to Tikonko with available data are Bo, Kenema, and Pujehun.

As an additional measure, it was decided that a regional analysis would be completed as well. This will allow for an additional point of comparison as the known on-site data will now be compared against the results of a geographic triangulation as well as a regional composition, potentially highlighting a rural community's ties to a regional identity. As Tikonko is located in the Southern Province, the list of 15 cities with available data was revisited, and all the cities located in the Southern Province were noted. The cities with available data that are also located in the Southern Province are Bo, Pujehun, Bonthe, and Moyamba. This selection of cities is displated in Figure 3.2.

Figure 3.2: A map of Sierra Leone showcasing the provincial borders divided into the Western Area, and Northen, Eastern, and Southern Provinces (Nations Online, 2014). As Tikonko, the focus of this case study, is located in the Southern Province, the attention in this map is drawn to the Southern Province as well. Circled in red is rural Tikonko while boxed in blue are the four cities in the Southern Province with data available in the Sierra Leone data portal of the Africa Information Highway. These four cities are Bo, Pujehun, Bonthe, and Moyamba (Africa Information Highway, 2020a).

Next, the data was filtered by "Source". The national statistical office of Sierra Leone, Statistics Sierra Leone (Statistics Sierra Leone, 2020), was selected as the other available data sources were almost entirely global and do not present data at a scale smaller than country-level. The process was repeated where the filtered data was examined to confirm urban-level granularity, and flagged accordingly. This resulted in a total of 13 datasets of varying topics all sourced from Statistics Sierra Leone. While all the relevant data can be fully visualized through the AIH interface, for the purposes of creating this methodology each dataset was downloaded and recompiled.

3.2.3 Determining Contextually Informative Indicators

To determine which specific indicators of the available datasets would be contextually informative, a few different processes were attempted until an optimal solution was determined. At first, the indicators present in Witmer's Predictive Tool survey (Witmer, 2018b) were utilized as a guiding set of questions. The Predictive Tool contains a list of 41 questions that help guide an on-site technical designer to make contextual observations. For example, the tool asks the technical designer to consider questions such as "Is there refrigeration in residences?" and "Do residents correlate illness with water-borne disease?". The initial process was to take this list of questions and attempt to answer each one using the remotely available data. However, while a suitable answer to questions similar to the refrigeration question were able to be found in the data, determining answers to a question based on opinion, such as the water-borne disease question, was much more challenging.

However, there is a secondary element of Witmer's Predictive Tool that is used to calibrate the on-site survey results with the national expectations of the country at hand. To generate national scores, Witmer curated a set of indicators sourced from global data banks such as the World Bank and the Food and Agriculture Organization. This set of 25 indicators was compiled by Witmer to be both contextually informative, as well as remotely available. As such, it was a much more appropriate frame of reference to guide the process of determining which indicators would be useful from the case study dataset.

In order to determine which of the available indicators from the case study data would be the most useful in predicting contextual conditions, over multiple iterations each of the 13 datasets were carefully parsed through to narrow down the selection. Of the 876 unique indicators from the available case study data, this iterative search for contextual informing indicators shrunk the selection down to 74 key indicator subjects. The process detailed in Section 3.2.4 will pull from this pool of 74 selected indicators.

3.2.4 Matching Indicators with Available Data

The 25 indicators Witmer selected for her calibration analysis are indicators that were collected uniformly across the globe, for nearly every country. These indicators are represented with 25 unique variables classified into five categories each represented by one of the five critical influences, cultural (abbreviated CU), political (abbreviated PO), economic (abbreviated EC), educational (abbreviated ED), and mechanical, formally technical, (abbreviated TE). The sources for each of these original indicators can be found in Appendix B. While the Predictive Tool makes sure to pull the most recent result for the country in question, it does not have to perform any modification process to ensure the data fits the tool, as the indicators are the same country to country. However, as one can imagine, the exact same indicators with the exact same wording were not consistently present in the data available for this case study. While some indicators had a near one-to-one match, for example Variable ED1, most indicators required some informed thinking to find appropriate counterparts.

If a near one-to one match could not be found, the following next steps were taken. First, an attempt was made to find a couple of data points from the case study data that could be used in combination to create a near match (see variable CU3). If this was not possible, the original indicators were reviewed to determine if a consolidation would be possible, using one indicator value to inform two variables (see variable CU2). If this was unsuccessful, the lack of comparable data was considered on its own merit, and potential conclusions were drawn (see variable PO4). Finally, if all other options were exhausted and unsatisfactory, the original indicator value as determined by Witmer's national level data was applied as the data point for all cities in the case study analysis (see variable PO5).

Once these matches are completed, a Likert scale was created for each of the new representative indicators. Witmer's original Likert scale for each indicator was not used for two reasons. First, as much of the data was not one-to-one comparable, the existing Likert scale did not measure the same values as the new data. Secondly, creating a new scale for each indicator allowed the city values to be measured against the country average. This allows the methodology to imitate the calibration procedure outlined in Witmer's Predictive Tool process. Each generated Likert scale for all of the case study data placed the country level average squarely in the middle of a one to five scale. Therefore the score assigned to each city is in relation to the national average, as opposed to a global standard. This process allows for a richer analysis, as the scoring highlights traits or characteristics of each city that stand out within the context of their own country. The ranges determined for each Likert scale were designated to ensure city level differences were effectively evident, while maintaining a reasonable incrementation amount. A sample of this process can be found in Figure 3.3 displaying the Likert scale and resulting scores for each examined city for variable ED1, the literacy rate for adult males.

Variable ED1

Sierra Leone	Literacy Rate, Male	49%
Bo	Literacy Rate, Male	74%
Kenema	Literacy Rate, Male	67%
Pujehun	Literacy Rate, Male	36%
Bonthe	Literacy Rate, Male	65%
Moyamba	Literacy Rate, Male	45%

Sierra Leone 49% set as middle of scale	
Determined range of available data	Score
< 34%	1
< 44%	2
< 54%	3
< 64%	4
else	5

Sierra Leone	ED1 Score =	3
Bo	ED1 Score =	5
Kenema	ED1 Score =	5
Pujehun	ED1 Score =	2
Bonthe	ED1 Score =	5
Moyamba	ED1 Score =	3

Figure 3.3: Exemplified is the scoring process which takes the raw data, in this case literacy rate for adult males represented by variable ED1, and translates it through a developed Likert scale to generate a score for each examined city. The scale is centered around the national average for Sierra Leone in each variable case. Data sourced from Africa Information Highway, 2020a.

Recalling the stated initial conditions, the following matches were conducted mimicking the decisions and connections a technical designer fairly new to the practice of contextual engineering may make, as this methodology aims to examine the potential for contextual engineering to be conducted remotely by those without formal or in-depth education in the discipline. While there is a reasonable justification for each match and connection, this is not to imply contextual agreement. Chapter 4 will spend time breaking down these decisions, analyzing not just the results of the methodology, but the contextual efficacy of these actions. A summary of these matches can be found in Table 3.1.

For reference, a higher cultural score indicates a greater significance of culture in the community. A higher economic score illustrates a community with more impactful financial limitations. A higher educational score demonstrates a greater importance of education to community members. A higher political score suggests a greater degree of political oversight and influence. Finally, a higher mechanical (formally technical) score indicates a greater degree of mechanical aptitude and innovative self-sufficiency to be present in the community. In this case study, all scores are with relation to the national average.

Table 3.1: Summary of variable abbreviations, the associated contextually-informing indicators determined by Witmer (Witmer, 2018b), and the matching indicator(s) that were selected from the Sierra Leone data portal of the Africa Information Highway (Africa Information Highway, 2020a) to most closely address the intent of the original indicators.

Variable	Indicator Name	Indicator Match (AIH)
CU2	CPIA policies for social inclusion/equity cluster average (1=low to 6=high)	Literacy Rate gender ratio, Parity Index (GCR) (Primary), Parity Index (GER) (SSS), Employment to Population gender ratio, Life Expectancy at Birth gender ratio
CU3	CPIA gender equality rating (1=low to 6=high)	Literacy Rate gender ratio, Parity Index (GCR) (Primary), Parity Index (GER) (SSS), Employment to Population gender ratio, Life Expectancy at Birth gender ratio
CU4	Bribery incidence (% of firms experiencing at least one bribe payment request)	*No available AIH match, national data used*
CU5	CPIA social protection rating (1=low to 6=high)	Number of Employees in Social Benefit Careers, Poverty Severity Index (P2)
CU6	Agricultural land (% of land area)	Distribution of Labour Force by Industry: % Agriculture, hunting and forestry
EC1	Adjusted net national income per capita (current US$)	Poverty Severity Index (P2)
EC2	Maternal mortality ratio (modeled estimate, per 100,000 live births)	Under-five Mortality Rate (U5MR)
ED2	Literacy rate, adult female (% of females ages 15 and above)	Literacy Rate Female
ED1	Literacy rate, adult male (% of males ages 15 and above)	Literacy Rate Male
ED4	Primary completion rate, female (% of relevant age group)	Primary Level Gross Completion Rate Female
ED3	Primary completion rate, male (% of relevant age group)	Primary Level Gross Completion Rate Male

Table 3.1 (cont.)

Variable	Indicator Name	Indicator Match (AIH)
ED6	School enrollment, tertiary (% gross)	Gross Enrolment Ratio (SSS)
ED5	School enrollment, secondary (% gross)	Gross Enrolment Ratio (JSS)
PO1	CPIA property rights and rule-based governance rating (1=low to 6=high)	Number of Employees in Property and Finance Careers
PO2	CPIA public sector management and institutions cluster average (1=low to 6=high)	Number of Employees in Public Sector Careers
TE1	Access to electricity (% of population)	Households by principal source of lighting: Electricity (NPA/BKPS)
PO3	CPIA quality of public administration rating (1=low to 6=high)	Number of Employees in Public Administration Career
PO4	CPIA transparency, accountability, and corruption in the public sector rating (1=low to 6=high)	*No available data on transparency, accountability, corruption, or other similar information. Total lack of transparency implied*
TE3	Fixed broadband subscriptions (per 100 people)	Households by main source of information: Internet Access (Social Media)
TE4	Renewable electricity output (% of total electricity output)	Households by principal source of lighting: Renewable (Solar)
EC3	Inflation, consumer prices (annual %)	Food Poor (%), Total Poor (%), Incidence of poverty (P0), Intensity of poverty (P1), Severity of poverty (P1)
PO5	Government Effectiveness: Estimate	*No available AIH match, national data used*
PO6	Political Stability and Absence of Violence/Terrorism: Estimate	*No available AIH match, national data used*

Table 3.1 (cont.)

Variable	Indicator Name	Indicator Match (AIH)
TE2	Improved sanitation facilities (% of population with access)	Households by principal source of drinking water: Piped or protected well
CU1	Religious Composition by Country 2010-2050	Households by main source of information: Religion

3.2.4.1 Variable CU1: Religious conditions

The original indicator measures the ratio of the number of members of the most prevalent religion against the sum of all religious people in the country. The most effective comparable indicator was determined to be the percentages of households that received their primary information from a place of worship. A higher percentage would indicate a larger number of dedicated religious members in that community, increasing the cultural score.

3.2.4.2 Variable CU2: Country Policy and Institutional Assessment (CPIA) policies for social inclusion/equity cluster average

The original indicator assesses policy for social inclusion including gender equality and equity of public resources. Determining a comparable indicator from the available case study data for this indicator was very challenging. As one of the key factors of this indicator is gender equality, variable CU2 was set to equal the determined value for variable CU3.

3.2.4.3 Variable CU3: CPIA gender equality rating

The original indicator examines the institutional systems within a country that inform gender equality or inequality. From the available data, four different indicators were selected to work in tandem to create the most accurate possible picture of gender equality. These four indicators were gender disaggregated literacy rate (converted to a female to male ratio), gender parity index (measuring female to male ratio), gender disaggregated employment to population ratio (converted to a female to male ratio), and gender disaggregated life expectancy at birth (converted to a female to male ratio). With regards to all four indicators, the closer a score is to one, the more equal the female to male split. The more equal a ratio is found to be, the lower the cultural score. These four scores were then averaged together to construct the value for variable CU3.

3.2.4.4 Variable CU4: Bribery incidence

The original indicator measures the percentage of public infrastructure firms that have experienced at least one bribery request. As there was no information about bribery or anything similar such as corruption in the case study data, the original national value was applied for all cities.

3.2.4.5 Variable CU5: CPIA social protection rating

The original indicator assesses the government policies and social protection programs that reduce poverty and ensure welfare. A combination of two indicators from the case study data were utilized to inform this variable. The first was a measure of the percentage of employees in each city that worked in a social protection career. These careers were selected to be Administration and support service activities, Public administration and defence compulsory social security, Education, and Human health and social work activities. The assumed correlation is the greater the percentage of social protection employees, the higher priority that service is to the community, and the greater their cultural score. The second indicator pulled from the case study data was the poverty severity index. As this index increased, the cultural score decreased. These two scores were then averaged together to construct the value for variable CU5.

3.2.4.6 Variable CU6: Percentage agricultural land

The original indicator examines the percentage of land area dedicated to agriculture. From the case study data, a similar objective was accomplished using the percentage of the labor force that was agricultural in nature. As the percent distribution of the agricultural industry increased, so did the cultural score.

3.2.4.7 Variable EC1: Adjusted net national income per capita

The original indicator measures the country's average income to determine the anticipated level of economic need. This goal was accomplished through the case study data by looking at the poverty severity index. As the index increased, the assumed economic need increased as well.

3.2.4.8 Variable EC2: Maternal mortality ratio

The original indicator measured the number of maternal deaths per 100,000 live births. While this exact statistic was not available, the under-five mortality rate was documented in the case study data. The greater the mortality rate, the higher the economic score.

3.2.4.9 Variable EC3: Inflation, consumer prices

The original indicator measured the percentage change in the cost of acquiring basic goods from year to year. As inflation data was not present in the case study data, a combination of poverty indicators and their change over time was utilized instead. These indicators included food poor percentage, total poor percentage, incidence of poverty, intensity of poverty, and severity of poverty. Positive correlation was established between growing poverty rates and higher economic need.

3.2.4.10 Variable ED1: Literacy rate, adult male

Adult male literacy rates were available from the case study data, no alternatives were necessary. As the literacy rate increased, the educational score increased as well.

3.2.4.11 Variable ED2: Literacy rate, adult female

Adult female literacy rates were available from the case study data, no alternatives were necessary. As the literacy rate increased, the educational score increased as well.

3.2.4.12 Variable ED3: Primary completion rate, male

Male primary completion rates were available from the case study data, no alternatives were necessary. As the completion rate increased, the educational score increased as well.

3.2.4.13 Variable ED4: Primary completion rate, female

Female primary completion rates were available from the case study data, no alternatives were necessary. As the completion rate increased, the educational score increased as well.

3.2.4.14 Variable ED5: School enrollment, secondary

The original indicator measured percent gross of tertiary school enrollment. The available case study data broke post-primary schooling into two categories, junior secondary and senior secondary. The indicator selected to fulfill this variable was the gross enrollment ratio for junior secondary school. As the enrollment ratio increased, so did the educational score.

3.2.4.15 Variable ED6: School enrollment, tertiary

The original indicator measured percent gross of secondary school enrollment. As stated, the available case study data broke post-primary schooling into two categories, junior secondary and senior secondary. The indicator selected to fulfill this variable was the gross enrollment ratio for senior secondary school. As the enrollment ratio increased, so did the educational score.

3.2.4.16 Variable PO1: CPIA property rights and rule-based governance rating

The original indicator assessed the presence and effectiveness of property and contract rights. This was mimicked by examining the percentage of employees in relevant career fields, including Financial and insurance activities, and Real estate activities. The conjecture is made that as the percentage of employees in these fields increases, the effectiveness increases, as does political oversight and therefore the political score.

3.2.4.17 Variable PO2: CPIA public sector management and institutions cluster average

The original indicator examined the quality of public administration. This was measured similarly by pulling the percentage of total employees in public sector careers. These careers included Administration and support service activities, Public administration and defence compulsory social security, and Human health and social work activities. The greater the percentage of employees in these careers, the higher the political score.

3.2.4.18 Variable PO3: CPIA quality of public administration rating

The original indicator measured the extent to which public services were effectively delivered to community members. This measurement was reproduced using the percentage of people in Public administration and defence compulsory social security, with positive correlation between this percentage and the political score.

3.2.4.19 Variable PO4: CPIA transparency, accountability, and corruption in the public sector rating

After parsing through all of the available case study data, it was determined that the datasets contained no information about transparency, accountability, corruption, or any similarly informing indicators. As such, the assumption was made that there was no such enforcement of transparency within the country. A total lack of transparency correlates with the highest possible political score.

3.2.4.20 Variable PO5: Government effectiveness

The original indicator measured the perceptions of the quality of public services and magnitude of political pressures. As this was a perception-based indicator, finding a meaningful alternative was not successful. Therefore, the original national value was applied for all cities.

3.2.4.21 Variable PO6: Political stability and absence of violence/terrorism

The original indicator measured the perceptions of the likelihood of political conflict or instability as well as terrorism. Again, as this was a perception-based indicator, no meaningful alternative was found in the case study data and the original national value was applied for all cities.

3.2.4.22 Variable TE1: Access to electricity

The original indicator evaluated the percentage of the population with access to electricity. A similar statistic was determined by examining the percentage of households that used electricity as their principal source of lighting. As the percentage increases, so does the assumed mechanical aptitude, and therefore the mechanical score increases as well.

3.2.4.23 Variable TE2: Access to sanitation facilities

The original indicator measured the percentage of the population with access to sanitation facilities. Using the case study data, a similar measure was evaluated by computing the percentage of households with access to plumbing or protected water sources, including piped indoor plumbing, piped in compound plumbing, and/or protected wells. As the access and use of these options increases, so does the mechanical score.

3.2.4.24 Variable TE3: Fixed broadband subscriptions

The original indicator examined the percentage of broadband internet subscriptions. To obtain a similar measure, the percentage of people that used the internet via social media as their primary source of information was pulled from the case study data. The assumption being that there is a positive correlation between the percentage of people who use the internet for their primary source of information, and the percentage of people who subscribe to the internet in general. As this value increases, the assumed technical aptitude decreases, as there is a lesser reliance on innovative rural self-sufficiency. Therefore the mechanical score decreases as well.

3.2.4.25 Variable TE4: Renewable electricity output

The original indicator observed the percentage of electricity generated by renewable sources. To mimic the information this indicator sought after, the percentage of people who used renewable energy sources (in this case solar) as their principal source of lighting was examined. As this percentage increases, so does the mechanical score.

The full calculations completed for each of the listed variables are available in Appendix C.

3.2.5 Calculating City Scores

Once all matches were completed, the scores determined for each variable were compiled into the appropriate critical influence equation (Equations 3.1 through 3.5), organized by city. These equations are the same ones determined by Witmer in her Predictive Tool and include variable weighting to more accurately represent the contextual impact each individual variable has on the overall critical influence score. Each city ultimately receives a score for each influence: cultural, economic, educational, political, and mechanical. To determine an influence's score for a particular city, the relevant variables are multiplied by their corresponding weights, summed together, then divided by the total sum of the weights, as represented in equations 3.1 through 3.5. After this is completed for each influence, the five scores are summed together and each influence is divided by that total sum, then each is multiplied by 100. This results in five proportional scores for each city. The higher the score, the more meaningful that particular influence is to the community in question.

$$Cultural = \frac{CU1 * 2 + CU2 * 2 + CU3 * 2 + CU4 * 2 + CU5 * 1 + CU6 * 1}{10} \tag{3.1}$$

$$Economic = \frac{EC1 * 2 + EC2 * 1 + EC3 * 2}{5} \tag{3.2}$$

$$Educational = \frac{ED1 * 1 + ED2 * 2 + ED3 * 1 + ED4 * 2 + ED5 * 2 + ED6 * 2}{10} \tag{3.3}$$

$$Political = \frac{PO1 * 1 + PO2 * 1 + PO3 * 1 + PO4 * 1 + PO5 * 2 + PO6 * 2}{8} \tag{3.4}$$

$$Technical = \frac{TE1 * 1 + TE2 * 2 + TE3 * 2 + TE4 * 1}{6} \tag{3.5}$$

The scores generated for each of the examined cities, Bo, Kenema, Pujehun, Bonthe, and Moyamba, as well as the national level scores, are visible in Table 3.2. As a reminder, the national level scores represented in the Sierra Leone column, are not values generated by this methodology, but instead are values calculated using the standard data for Sierra Leone with the unmodified Predictive Tool. In other words, these are the anticipated average country-level scores for Sierra Leone. The greatest value, and therefore most impactful, of all the influences is bolded for each column.

Table 3.2: The five critical influence scores calculated for each examined city, Bo, Kenema, Pujehun, Bonthe, and Moyamba, as well as the national scores for Sierra Leone. The highest value, representing the most impactful influence, is bolded in each column.

	Sierra Leone	Bo	Kenema	Pujehun	Bonthe	Moyamba
Cultural	25.4	14.8	18.1	24.7	21.5	22.5
Economic	**33.3**	19.5	18.5	**30.2**	13.1	**28.1**
Educational	8.7	**30.4**	**28.8**	13.7	**32.0**	14.7
Political	18.1	16.0	14.4	15.4	16.4	13.6
Technical	14.5	19.3	20.2	16.0	17.0	21.3

With all of the scores for each city compiled, city weighting was applied in order to attempt a more accurate triangulation process. Equation 3.6 below displays the weighting process applied for the cultural critical influence. The same equation, with the appropriate variable scores substituted, was utilized to calculate the weighted value of the remaining four critical influences as well. The closer a city is by geographic distance to Tikonko, the stronger its influence will be on the calculated score. Table 3.3 provides further information on the specific distances used in the weighting calculation as well as travel time estimates for three different methods of transportation, driving, cycling, and walking, as determined by OpenStreetMap (OpenStreetMap contributors, 2021). Important to note, Bonthe is accessible only by ferry. OpenStreetMap cannot estimate a time directly from Tikonko to Bonthe. Thus, the values displayed for Bonthe are calculating the driving distance from Tikonko to coastal Yargoi, Sierra Leone, while the travel times all have been increased by two hours to account for the time spent on the transport ferry (VSL Travel, 2021). This travel time information is not factored into the quantitative weighting analysis, but provides insight as to the estimated level of interaction the two communities may have.

$$Cultural\ Weighted = \frac{\sum(Cultural_{city\ i} * \frac{1}{Distance_{city\ i}})}{\sum \frac{1}{Distance_{city\ i}}} \tag{3.6}$$

Table 3.3: The measured distances in kilometers between Tikonko and each urban city in question and the estimated travel times over three different methods; driving, cycling, and walking (OpenStreetMap contributors, 2021).

	Bo	Kenema	Pujehun	Bonthe	Moyamba
Distance from Tikonko (km)	12	81	84	103	117
Travel Time: Driving (hours)	0.32	1.10	1.63	4.82	1.90
Travel Time: Cycling (hours)	0.98	5.97	4.80	9.72	7.95
Travel Time: Walking (hours)	2.37	16.52	15.33	23.12	19.87

Two weighted estimations are performed. The first is a triangulation between the three cities closest to Tikonko by kilometer distance. These cities are Bo, Kenema, and Pujehun and are referenced in Table 3.4 as the Tikonko triangulation estimate. The second estimation examines only cities located in the same province as Tikonko, Sierra Leone's Southern Province. These cities are Bo, Pujeun, Bonthe, and Moyamba and are referenced in Table 3.4 as the Tikonko Province estimate. As in Table 3.2, the greatest influence value of each estimation is bolded. In both cases presented in Table 3.4, the bolded value belongs to the educational influence.

Moving forward, the critical influence scores computed through this methodology and displayed in Table 3.4 will be referred to as the scores for cultural, economic, educational, political, and technical, as opposed to the weighted version of each

Table 3.4: Weighted critical influence scores calculated for Tikonko using two different sets of cities. First, the Tikonko triangulation estimate, a triangulation of the three closest cities to Tikonko, Bo, Kenema, and Pujehun. And second, a Tikonko province estimate, using all cities with available data in the Southern Province, Bo, Pujehun, Bonthe, and Kenema.

	Tikonko Triangulation Estimate	Tikonko Province Estimate
Cultural Weighted	16.3	17.0
Economic Weighted	20.5	20.7
Educational Weighted	**28.4**	**27.6**
Political Weighted	15.7	15.8
Technical Weighted	19.0	18.9

influence. This is to ease the comparison process reviewed in Chapter 4.

CHAPTER 4

RESULTS AND DISCUSSION

This chapter discusses the data sources considered, the structure for global applicability, and the efficacy of the methodology proposed in the previous chapter. Efficacy is evaluated by examining the results of the case study and providing a summary of some of the challenges and problems observed during the process.

4.1 Data Source and Acquisition

The nature of this work was to create an accurate contextual representation of a specified rural community using solely remotely accessible data. There exist a variety of data banks and sources that were explored to be used in this research, each with benefits and shortcomings. The World Bank, for example, provides mostly uniform data across the globe, is collected from a relatively reliable source, and is updated with regularity. However, the smallest scale of data available from the World Bank is at the country level (World Bank, n.d.) which is decidedly not local enough to gather contextual data.

On the other side of the spectrum, there exist smatterings of local data gathered minutely with great attention to detail and emphasis on social objectives, similar to those highlighted by contextual engineering. However, as one can imagine, this data is few and far between. It is hardly available for every rural community in question, and in the case that the data is available, it is often contained within a thorough report or analysis seeking to accomplish an established purpose, as opposed to objective data collecting for the sake of future use in any variety of ways.

The final major consideration to be accounted for was accessibility of the data. The hope of this research was that a methodology could be created using remote data inputs to return a contextual understanding output for nearly any rural community of interest. For this goal to be accomplished, the data utilized must be similar enough between locations for the methodology to be applicable at a global scale.

With these considerations in mind, it was determined that the most appropriate data source for this methodology was national statistical office data. The data collected by the country of origin is commonly available at the urban level as opposed to the national, collected for record keeping purposes as opposed to a specific agenda, and is consistent enough across the globe to be workable, checking off all of the desired boxes. However, national statistical office data does pose its own obstacles. One such obstacle is that each nation will collect the unique information it feels is relevant and thus data banks will not be one-to-one comparable. Additionally, as each national databank collection is unique, it is not possible to predict or account for all of these obstacles in advance. These challenges are elaborated on further in Section 4.3.3. However, the overwhelming benefit of using national statistical office data was the ability to view disaggregated data at the urban level. It was thus hypothesized that this national statistical office data, available for nearly all nations and collected at this relatively micro scale, was the best available data for creating a generalized methodology allowing technical designers to understand local context from abroad.

As this research operated under the realistic premise of time and user experience limitations, it aimed to explore a methodology appropriate for these circumstances. It was not attempting or expecting to fully replace the experience of on the ground community interaction, but rather examine the result of an technical designer undertaking the valuable task of collecting contextual data, without the resources to do so face-to-face or the educational background to make certain informed leaps that may be necessary during the data collection and analysis process. The value of this research was exploring the outcome of a unique and previously unstudied application of contextual engineering that is rooted in need and a progressing reliance on big data.

The national statistical office data utilized by a technical designer for their project in question will generally be available in a variety of formats. Through evaluation of a variety of these resources, this research has created guided recommendations. Integrated Public Use Microdata Series (IPUMS) is one database that was evaluated for its effectiveness within this research context. IPUMS is an organization which collaborates with over 100 national statistics offices to offer a database of census microdata (IPUMS, 2019). As such, each record in this database represents one person. While this has potential to be useful as a supplementary data set, as mentioned earlier, it does not lend itself to an application of statistical urban-level data. Additionally, technical designers intending to use this data must first submit an application for an account which will be reviewed before access can be granted. The added steps required to access this data are counteractive to the intent of a methodology that is appealing to a user operating under possible time and/or monetary constraints. While this research hopes the case of incorporating contextual data has been sufficiently made, it is important to consider that its application is dependent on the interest level of the technical designer. Thus, it is important the data proposed for use is relatively easy to access.

Locating the data at the source, for example, the United States Census Bureau (United States Census, 2020) site for data about the United States, or the Statistics Sierra Leone (Statistics Sierra Leone, 2020) site for data about Sierra Leone was another potential avenue. However, there are a variety of challenges associated with this option. For example, not all statistics offices use English as their primary language. This results in a reliance on web translation tools which do not always provide accurate translations, and cannot translate all elements of the site including images or downloadable content. Another challenge was with regards to relaying the methodology. It would not have been feasible for the scope of this work to create and recommend a methodology that accounts for any variety of potential source formats, so cite-specific source data was not specifically recommended. While national statistical office data located at the source is workable, a more refined method was sought after for this research.

Knoema (Knoema, 2020) is an online tool that was evaluated to address the previously stated concern of lack of uniformity among data. While the data itself remains in its original form, Knoema collects data of nearly 200 different countries and categorizes this data to be accessed through an interface that is largely consistent between countries. While this addressed the search for a comprehensive tool, Knoema is unfortunately not free to fully access. As this research hypothesizes that cost may be a constraint for project technical designers, suggesting a paid service is not aligned with this objective.

A partial solution to the issue of cost was found through a database created by the African Development Bank called the Africa Information Highway (AIH). Launched in 2012 as a part of the Statistical Capacity Building Initiative program (African Development Bank, 2020), the AIH aims to be globally beneficial, supporting both the improvement of data collection in African countries and the public distribution of this data internationally. As such, the data available on the AIH is a collection of 54 open data portals, one for each African nation. The Open Data Platform utilized is "based on the Knoema IT platform" (African Development Bank, 2017), mentioned above. The data available is nearly identical to the data found through Knoema, as the AIH similarly gathers and disseminates data from a variety of sources including the census data of countries, as well as global data banks such as World Bank, Food and Agriculture Organization, and United Nations agencies. However, an important distinction is that data accessed through the AIH is free for all users. And importantly for this geographically concerned research, this interface allows users to sort data by topic, source, or region/province/state (as applicable). By selecting the geographic category or the source category, the user are able to view the geographically disaggregated data or all data sourced from the national statistical office, respectively.

While using the AIH database for this research was limiting to only projects within African nations, the data available for nearly 150 other global nations is available in a very similar format through Knoema. This methodology is relatively applicable for any global community, but the lack of open data is a recognized obstacle. Further discussion on encountered problems and obstacles can be found in Section 4.3.3.

4.2 Global Applicability

While the methodology outlined in Chapter 3 was conducted with specific relation to the community of Tikonko, Sierra Leone, the process was designed to be flexible to adoption for many rural communities across the globe. The efficacy of this approach will be examined further in this chapter, but the following describes the intended adjustments necessary for global applicability.

There are a few initial conditions that must be met in order for a rural community to be analyzed in a similar way to methods outlined in this research. First, the community must be located within a country recognized by either of the two recommended data sources, Knoema or Africa Information Highway. If the country is not recognized within AIH but is available within Knoema, funds must be procured to access the relevant country's database, as Knoema is not open source.

Next, the available data for the country in question must be examined for the smallest granularity documented, as some countries do not formally collect data at the urban level. If urban-level data is available, the three cities from the available data in closest geographic proximity to the rural community in question should be noted. At this point, the geographic distance between the rural village and each city should be documented as well.

The next step would be to review the list of 25 indicators determined by Witmer and parse through the available data to find the most appropriate match for each, following the same process of the case study methodology. Once the indicators are all matched, Likert scales are created, and scores are assigned, equations 3.1 through 3.5 can be utilized to generate critical influence scores for each city. Finally, the critical influences for each city are combined and weighted by distance using equation 3.6 appropriately for each of the five influences. This generates the estimated community score where the greatest value informs the technical designer of the community's most dominant influence.

While this would be the procedure followed by a practitioner attempting to utilize this methodology for a rural community other than Tikonko, Sierra Leone, there are some observed challenges that inhibit the author from recommending this process fully. The challenges are discussed in the following sections, Section 4.3 and Section 4.4.

4.3 Case Study Analysis

This section outlines the on-site data collection process with resulting scores and compares them against the scores calculated by the methodology. The efficacy of the methodology is evaluated and challenges with the case study data are discussed.

4.3.1 On-Site Tikonko Data Collection

In the summer of 2019, the author, Witmer, and a third contextual engineer traveled to Sierra Leone to visit Bo, Tikonko, and various smaller villages throughout the Tikonko Chiefdom. The intent of the trip was not for any formal project scoping or future implementation, but rather solely intended as an opportunity to learn from a variety of people and communities. The information gathered and connections made on this trip has formed and will continue to form the foundation for educational project partnerships and programs in the years to come. As such, this team spent a week in Tikonko proper conducting interviews, taking notes, and interacting with as many members of the community as possible. Using the knowledge gleaned from these interactions and observations, the Predictive Tool was completed by each of the three team members for the village of Tikonko. Table 4.1 below displays these Predictive Tool critical influence score results. Bolded in each column is the highest value calculated by each team member's results, coordinating with the most dominant critical influence determined.

Table 4.1: Computed critical influence scores for Tikonko, Sierra Leone as tallied from the results of the Predictive Tool survey completed by each of the three team members during the 2019 travel. The greatest value in each column is bolded, and corresponds with the community's most dominant critical influence.

	Team Member A	Team Member B	Team Member C	Team Member Average
Cultural	15.7	21.0	21.0	19.2
Economic	19.8	16.6	20.6	19.0
Education	**23.7**	**23.3**	**22.4**	**23.1**
Political	20.5	19.2	17.7	19.1
Technical	20.4	19.8	18.4	19.5

As a clarifying note, the scores presented in Table 4.1 are not exact matches between each team member, even while the team members are all observing the same community. This variation is expected and a result of the unique experiences and interpretations of observations that each individual will make, thereby influencing their responses to identical survey prompts. In addition, while the team members were together for the vast majority of the trip, there were times where the group separated, resulting in some uniquely experienced moments, events, and conversations. The combination of these two known and anticipated factors resulted in variation between the calculated critical influence scores. However, in accordance with Witmer's analysis process and methodology, the results are valid as long as the highest scoring influence for each team member is the same. As this is the case with the on-site scores presented in Table 4.1, the data is valid and an average of the scores can be computed and utilized as the prevailing result.

The resulting dominant influence for the community of Tikonko was found to be education. As previously detailed, a dominant influence of education corresponds to a community that highly values the educational process and seeks knowledge eagerly. Having this understanding of a community's predisposition to learn is crucial to a technical designer as it influences each step of the project process. For example, a community such as Tikonko with a prevailing educational score is likely to want to provide much more input on the project design, and receive much more thorough status reports than a community without a dominating educational trait. Incorporating this understanding throughout the project process has been found to result in an increased rate of project success (Witmer, 2018b).

4.3.2 Comparison of On-Site Scores with Calculated Results

In both the Tikonko Triangulation Estimate and the Tikonko Province Estimate, the results generated from the methodology indicated Tikonko's prevailing critical influence to be education. This conclusion appears to be supported by the results of the on-site Predictive Tool surveys completed by the team of contextual engineers, represented by the column titled Team Member Average. Interestingly, the results computed by the unmodified Predictive Tool to calculate the country-level scores determined that the economic influence stood far ahead of the other critical influence. This would suggest that financial constraints are the primary influence dictating the design and implementation goals and limitations, contrary to the results of the on-site Team Member Average which demonstrates the educational influence reigning. At a base level, it would appear that the urban-level data provided a more successful approximation of the Tikonko conditions than a technical designer would achieve examining just the remotely-accessible country data from the unmodified Predictive Tool on its own. The direct comparison is observed in Table 4.2.

These results seem to indicate the methodology was successful in replicating the priority score determined by the on-site team. In this particular case, both the on-site investigation and remotely-accessed contextual data triangulation methodology would result in the conclusion that Tikonko is a community with an affinity for learning and thirst for knowledge. Both results would encourage a technical designer to structure their project process around the educationally dominant trait, the guidelines of which are elaborated on by Witmer's work. However, the intention of this research goes beyond generating a successful match

Table 4.2: Comparison of the country level anticipated average scores as determined by the unmodified Predictive Tool against critical influence scores computed from the methodology against the scores resulting from the on-site data and analysis. The highest value in each column is bolded and represents the most prominent influence. The Sierra Leone country-level column indicates the Economic influence to dominate while the others indicate the Education influence to dominate.

	Sierra Leone	Tikonko Triangulation Estimate	Tikonko Province Estimate	Team Member Average
Cultural	25.4	16.3	17.0	19.2
Economic	**33.3**	20.5	20.7	19.0
Education	8.7	**28.4**	**27.6**	**22.4**
Political	18.1	15.7	15.8	19.1
Technical	14.5	19.0	18.9	19.5

and instead intends to analyze the efficacy of the methodology process altogether and discuss problems observed with the data. Section 4.3.3 will elaborate on the challenges associated with the case study data in particular, while Section 4.4 will expand upon the difficulties and obstacles associated with this remote-based process as a whole.

4.3.3 Challenges with the Data

While the methodology appeared to result in critical influence scores that aligned with the actual on-site scores, there were a number of hurdles along the way that may have had an impact on the final result of the methodology. These challenges were unavoidable and unpredictable, and relied on the judgement of the author to overcome them throughout the process of the methodology. As such, they are crucial to discuss as it can be assumed that not every practitioner of the methodology would address the challenges in the same way. It is also worth noting that although the author adopted the persona of a technical designer not contextually familiar with the Tikonko community, lingering unintentional bias may have affected some of these decisions as well. This section will further outline observed challenges.

In order to find matches for the contextual indicators as described in the methodology, the practitioner must parse through a large quantity of data in order to determine which indicator or set of indicators is the most applicable, or determine that a match does not exist. There are many ways to logistically go about this process, but a good place to start is the use of keyword searches. For example, if the desired indicator is the male literacy rate, searching through the available data for the word "literacy" is a practical and likely effective option. However, this becomes much more difficult when the country of interest collects data in a language different to that of the practitioner. In the case of the Sierra Leone data, while the primary language utilized was English, the tendency was towards British English. This resulted in some spelling discrepancies between the expectation in the keyword search, and the actual presentation of the word. For example, the word "feces" was a keyword utilized in the search for a match for the access to sanitation facilities indicator, represented by variable TE2. However, the Sierra Leone data utilized the British English spelling of "faeces" which was not immediately apparent.

Along the same lines, spelling mistakes in the data were an encountered difficulty as well. For example, Figure 4.1 displays data from the Household Characteristics of Sierra Leone dataset in response to the Households by principal source of drinking water indicator. This table is pulled directly from the visualization feature of the

Sierra Leone Data Portal of the Africa Information Highway. Specifically, the figure is displaying a misspelling of the word "Neighbour's". In the official data, the spelling is listed as "Neigbour's". While this particular misspelling did not impact the keyword search process, it did present the concern that other misspellings may go unnoticed.

Region	Variable	Units	2015
Bo	**Households by principal source of drinking water**	Number	102,723
	Neigbour's Tap	Number	1,826
Pujehun	**Households by principal source of drinking water**	Number	51,514
	Neigbour's Tap	Number	79
Bonthe	**Households by principal source of drinking water**	Number	32,538
	Neigbour's Tap	Number	94
Moyamba	**Households by principal source of drinking water**	Number	61,880
	Neigbour's Tap	Number	742
Kenema	**Households by principal source of drinking water**	Number	111,734
	Neigbour's Tap	Number	5,964

Figure 4.1: Image of visualized data from the Sierra Leone Data Portal of the Africa Information Highway. The presented data is from the Household Characteristics of Sierra Leone dataset and is displaying an example of an encountered misspelling (Africa Information Highway, 2020a).

During the examination of another indicator from the case study data, it was observed that some of the presented numbers were not adding up. Figure 4.2 displays a sample of data available from the Employment and Labour Force of Sierra Leone dataset. Specifically, it calls out the employed population 10 years and above by industry and sex. Examples of two industries are shown in Figure 4.2, Public Administration and Defence Compulsory Social Security, and Education.

The Sex column is intending to display the total number of people working in the selected industry, as well as the separation of employees by Male and Female sex. However, as one can observe, none of the totals presented in Figure 4.2 are the sum of the relevant Male and Female values. In addition, in five of the six cases presented in the sample data the Male count is actually higher than the Total count altogether. While it is possible this is not an error with the data but a misunderstanding by the author, there is no other immediate explanation, nor does this discrepancy occur in any of the other instances of sex disaggregated data.

Location	Indicator	Sex	Units	2015
Bo	Public Administration and Defence Compulsory Social Security	Total	Number	1,423
		Male	Number	1,575
		Female	Number	507
	Education	Total	Number	2,082
		Male	Number	3,105
		Female	Number	1,511
Pujehun	Public Administration and Defence Compulsory Social Security	Total	Number	351
		Male	Number	698
		Female	Number	121
	Education	Total	Number	819
		Male	Number	1,291
		Female	Number	347
Kenema	Public Administration and Defence Compulsory Social Security	Total	Number	1,199
		Male	Number	1,444
		Female	Number	418
	Education	Total	Number	1,862
		Male	Number	1,831
		Female	Number	725

Figure 4.2: Image of visualized data from the Sierra Leone Data Portal of the Africa Information Highway. The presented data is from the Employment and Labour Force of Sierra Leone dataset and is displaying an example of an apparent math error (Africa Information Highway, 2020a).

When making indicator matches, in the case where there was no suitable counterpart available in the case study data, a logical assumption had to be made. In some cases, variables PO5 and PO6 for example, this resulted in falling back on the country-level score for that particular variable, as determined by the Predictive Tool. However, in one instance, the lack of existing data was evaluated on its own merit in order to draw a conclusion and inform the relevant indicator. This was the case for the rating of CPIA transparency, accountability, and corruption in the public sector, variable PO4. As there did not exist any information related to transparency, or anything even tangentially related to transparency, the conclusion was drawn that therefore Sierra Leone did not value transparency in the slightest, and the variable was scored accordingly. However, while Sierra Leone's corruption rates do not rate too favorably on a global scale, there is markedly low corruption or bribery at the local community level. In addition, Sierra Leone has notably made the reduction of remaining corruption a priority at the national level in recent years (Deutsche Welle, 2020). While we made the best assumption we could with the assumed knowledge we had acting as a student technical designer unfamiliar with Sierra Leone or contextual engineering, we were bound to make educationally-based leaps of logic that were ultimately inaccurate.

Another data challenge observation made was in relation to the impact some of the largest cities could have on the national average of an indicator. The Likert scales in the methodology were created to place the national average squarely in the middle of the one to five scale, and in most cases this appeared to be a suitable method. However, there was one notable case where the national average was intensely skewed by one urban area alone. Variable TE1, access to electricity was matched in the case study data with the percentage of houses utilizing electricity as their principal source of lighting. For this indicator, the national average was found to be 17.8%. Yet Bo, the second largest city in Sierra Leone was found to have an electricity use percentage of only 15.14%, two percentage points lower than the national average.

It was found that the reason for this discrepancy was the massive pull of the statistic connected to Freetown, the largest city in Sierra Leone and the country's capital. At 67% Freetown's electricity use percentage dwarfed the other measured values and single-handedly raised the national average to a value above every other city's.

Thinking as a contextual engineer, questions were raised about one's ability to interpret the case study data without a thorough understanding of the bigger picture, or the context in which certain statistics were observed or collected. While this hesitancy will be explored in greater detail in Section 4.4, some case study specific examples are worth noting as well.

Some indicators may suggest a certain understanding, but upon closer inspection may not mean much without further context. For example, employment titles listed in the Employment and Labour Force of Sierra Leone do contain associated metadata about the qualifications of the counted employees. Are employees only formally counted as Maternal and Child Health (MCH) Aides if they have a certain degree or certification? Are natural medicine practitioners in the same field counted as well? Are employees in more than one title counted multiple times? These are important questions to ask as they provide insight to a community's value system, but are not addressable with the data as it is presented.

A simple example of this occurrence was observed in the Household Characteristics of Sierra Leone dataset, specifically regarding the indicator determining households by main source of information. One of the response options to this indicator was "Church/Mosque". This was the only reference to religion in any capacity throughout all of the 13 examined datasets of the case study. There are a few conclusions one could draw from this, one being that as religion is hardly discussed, it is not of importance to the country. Another practitioner may interpret that any religious information being gathered at all in a formal statistical survey is an indication that religion is valued in the country, even with it being such a minor statistic. Another may observe that as church is listed before mosque, the country must be predominantly Christian and secondarily Muslim. All of these interpretations are rooted in logical judgement, yet there is no way to know who is correct just from an examination of the presented case study data. In this particular case, none of the interpretations were wholly correct. Sierra Leone is a country that is roughly three quarters Muslim and one quarter Christian (Pew Research Center, 2016) with a government that requires religious freedom in its constitution and a culture that supports common intermarriage between religions, sharing of holiday celebrations, and multiple religions living together in the same household (United States Department of State, 2019).

Another example of this contextual misdirection occurring in the Sierra Leone data is with regards to the Nuptiality and Fertility in Sierra Leone dataset. A practitioner examining this data may come to the conclusion that Sierra Leone values marriage at a young age. This perhaps could have cultural or religious implications, but seems to be the norm across the country, visible in the data presented in Figure 4.3. However, this particular dataset contained an associated executive summary from Statistics Sierra Leone stating some of their conclusions about the data they gathered. The Nuptiality dataset was the only one found to contain such a statement. In this statement, they call out this mean age at marriage statistic and note: "Sierra

Leone's Customary Marriage and Divorce Act states that girls cannot marry before 18 years of age. Yet the data suggests that the singulate mean age at marriage is less than 18 years across the country. This seems to indicate that the Act is not being implemented effectively." (Statistics Sierra Leone, 2015). This statement suggests that not only is the young marriage something the country is trying to fight against, it is actively not allowed. Associated with the contextual interpretation of this additional information is the consideration of the voice delivering the message. In this case, as the statement was released by an agency serving the government, it is possible there were largely political motivations at play as well. All of these factors add layers of interpretation to the data on its own, but are not readily available when examining this big data at face value.

Variable	Region	Units	1985	2004	2015
Marriage for the Female Population 15-54 Years, All	Sierra Leone	years	18.00	21.00	17.90
	Kenema	years	17.00	19.70	17.20
	Bo	years	17.60	20.50	17.80
	Pujehun	years	17.10	19.50	17.70
	Moyamba	years	17.80	19.30	16.60
	Bonthe	years	17.30	19.50	17.20

Figure 4.3: Image of visualized data from the Sierra Leone Data Portal of the Africa Information Highway. The presented data is from the Nuptiality and Fertility in Sierra Leone dataset and is displaying the Singulated Mean Age at Marriage for the Female Population 15-54 Years (Africa Information Highway, 2020a).

This section highlighted some of the recognized problems with the case study specific data. The following section will take a step back and examine some of the more fundamental challenges with using remote data to inform context.

4.4 Challenges with a Remote Data Process

This section will highlight many observed challenges with the remote-data process and methodology. This is not to say that the methodology was not a valuable exploration of the potential combination of big data and contextual engineering. Rather, this emphasizes the irreplaceable importance of thorough communication, in-person observation, and lived experience.

One foundational challenge with this process is that Witmer's contextual engineering survey is intended to be completed by on-site technical designers. As such, when the Predictive Tool intakes their survey responses and outputs a critical influence score for their community, there is an understanding that the technical designers will interpret the results through the lens of their lived on-site experience. The tool will guide them into considering the appropriate contextual elements that will impact their design and implementation process, and the lived experience will help provide the necessary and appropriate contextual details to make it happen. Without the lived experience hand-in-hand with the critical influence scores, the practitioner of this methodology is missing out on that contextual element. Of course, this was a known constraint from the start of the process, but it is important to keep in mind.

One of the key elements Witmer's work with rural communities emphasizes in contextual engineering design is the incorporation of rural innovative self-sufficiency. These are techniques and practices developed and grown within a community in response to obstacles or stresses. For example, Figure 4.4 displays an instance of rural innovative self-sufficiency in a small village in the Tikonko Chiefdom on the outskirts of Tikonko proper. The armrest of a plastic chair in the community had split, and was sutured back together with an organic material. This was not something that was intentionally shown to us, but an observation that was made while on-site. In contextual engineering applications, the observation of the repaired plastic chair would encourage technical designers to consider the willingness of community members to make repairs to existing items or infrastructure, the community's access to and use of spare parts and repair material, and other contextual considerations that influence decisions to be made during the design and implementation process. This type of observation is not something that could be indicated by the available remotely-accessible case study data.

Figure 4.4: Image of a rural innovative self-sufficient response to a broken piece of furniture in a small, rural village in the Tikonko Chiefdom of Sierra Leone. The armrest of the chair has been strongly sutured with organic material.

Distance weighting was selected as an ideal averaging method to ensure that the interpolation of a rural community was approximated based on an educated hypothesis. The postulation was made that the closer an urban city was to the rural community in question, the greater interaction those two communities were likely to have. This could be due to familial connections, trade or business deals, educational opportunities, etc. However, calling back to Table 3.3, some of these travel times were quite extensive, especially for community members that did not have a car or motor vehicle. At some point it is theorized that the distance becomes too great to make a worthwhile correlation, however this was not explored in depth in this particular study. Conducting a regional analysis was one selected method utilized as an alternative to a distance triangulation, but one may have more success with other methods such as only considering cities within a certain distance, or weighting cities with populations more similar to the rural area higher than those with much larger populations and forgoing a distance weight altogether.

As this methodology is interpolating rural conditions from urban data, it is necessary to consider the predictable differences between the two. One key difference is that urban areas tend to have broader access to resources than rural counterparts. This disparity particularly presents itself in healthcare and mental health services (Cherry et al., 2017), economic stability, affordability ratios of goods and services, and education (Bouck, 2004). There are some conditions that tend favorably to rural areas such as lower crime rates (Pateman, 2011), but for the most part, access to resources that improve well-being and the relevant correlating statistics are more likely to be found in urban areas. While this was not directly accounted for in the methodology, it is encouraged that any practitioners intending to use urban data to learn about a nearby rural community keep this consideration in mind as they interpret the available data .

When analyzing data that one did not explicitly collect, there are certain assumptions that must be made about the content of the data. This is especially true in this methodology as we use the available data to inform general community context, something the data was not necessarily initially collected to do. For example, while translating the available data to match the contextual indicators, the assumption was made that a greater Poverty Severity Index score correlates with a community that does not highly value social support systems for the mass population and therefore the community's cultural identity score would decrease. However, there are other potential conclusions that could be drawn about communities with a vast amount of poverty including purely economic reasons. Without knowing the broader context of the Poverty Severity Index score, it becomes very challenging to draw any formal conclusions with confidence.

While the original intention of this methodology was to create a process that could be molded to fit nearly any rural community desired, this proved to be much more challenging than anticipated. And as a result, drawing conclusions from the case study must be done with this uniqueness in mind. Making the process globally applicable is difficult for a number of reasons, the first being that every country collects different data. While the data available for Sierra Leone spanned thirteen unique datasets all disaggregated to the urban level, not all countries had such a wide breadth of data available. For example, Ghana only has two urban-level disaggregated datasets with only one of those two containing information on more than one indicator (Africa Information Highway, 2020b).

The language barrier is another important element to consider in the evaluation of the globalization of this process. While English is one of the national languages of Sierra Leone, and the national statistical office data was collected and presented in English, this is not the case for all countries. Relying on an online translation service adds one more barrier to understanding the original intention of the data, and has the potential to be misinterpreted.

Another significant consideration to keep in mind when utilizing big data to better understand a population, is that it can be easy to manipulate the data to end up with a favorable outcome. For the purposes of this methodology, the final critical influence scores were calculated only after all the variables were individually determined, and they were not altered after the fact. This was for transparency of communicating that none of the indicators were selected for the ability or lack thereof to align with the on-site scores. However, an ill-intentioned practitioner could disregard this transparency without too much trouble, which is a sizable concern.

On the reverse side, the data collector also has the potential to manipulate the collection or presentation of the data in a way that is favorable to their agenda, potentially even without malicious intent. For instance, a government may see potential to improve their global standing by manipulating their poverty statistics and be motivated to falsify the data presented. Collecting data yourself, as is done in contextual engineering, is one way to ensure the data collection process is transparent from start to finish, but this is not possible in this particular methodology.

Another initially unrealized obstacle with a full reliance on big data was the loss of the human element and connection behind the numbers. Sierra Leone, for example, has some of the world's highest maternal and under-five mortality ratios. As of 2019, over 10% of children in Sierra Leone will pass before they turn five years old (World Bank, 2019). When simply examining these numbers on a screen, one can forget the devastating real-world impact these values really have. Our travels to Sierra Leone were just a few years after the Ebola epidemic raged through the country. As we met with community members in various towns we were told harrowing tales from grieving family members and friends who watched as their loved ones were taken from them too soon. The impact hearing these stories first hand has not come close to being mimicked by reading about the experiences or looking at the data.

While the methodology did generate a result similar to the one produced by the on-site visit, the analysis examines how this alignment is not necessarily a vindication of the methodology's efficacy. A multitude of factors contributed to unpredictable obstacles along the way that may have unintentionally skewed the data, and the variability in the indicator interpretations leaves room for a number of different, yet potentially reasonable outcomes. However, the value of this methodology did not lay solely in its ability to accurately predict community conditions, but also in the exploration of the possibilities of remote data and their potential to aid in the understanding of context.

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

No population is monolithic, even within one city or community, and there is much that cannot be learned without directly interacting with the community. By extending the use of remotely-accessible data in this manner an immense number of assumptions must be made, beyond the comfort level of the author. But to reiterate, the challenges presented in the process of using remote-data to inform context do not detract from the value of exploring the limits of contextual research. There assuredly exists the need for an inclusion of more contextually-minded work as well as the need for adaptive solutions to travel restrictions and other limiting constraints.

Even if this research could not confidently provide answers to these needs, it can certainly act as a stepping stone and guide for future research. This includes the acknowledgement of some key takeaways to be considered for future use cases.

Remote data has limitations to its use. When comprehensive data was not available, assumptions and educated leaps of logic had to be made in order to fill the gaps. This resulted in interpretations that would likely be unique to each practitioner of the process. Making these less confident connections both weakens the conclusions that can be drawn, as well as introducing new questions and room for bias.

While contextual engineering does use similar data and indicators to those explored in the case study methodology to achieve an understanding of a community's context, the lived experience that accompanies the contextual engineering data collection process is irreplicable. This experience exposes decision-makers to elements of a community's make-up that simply cannot be captured through big data. Instead, if this process is to be repeated, the suggestion is made to include a face-to-face element into the analysis.

One potential option to introduce a face-to-face communication element is through the use of video chats. However, this poses its own limitations. For one, some stakeholders may not have reliable access to video chatting services. Language barriers may be an issue as well as coordination with a translator becomes necessary and communication over video chat is hindered as is. Additionally, there are still many elements that would be missed in comparison to an on-site visit including interaction with a plethora of community members and visual observations beyond the boundaries of the video call screen.

While remote-data on its own may not be able to provide sufficient context for understanding a community's conditions, it can certainly still provide some value. Decision-makers who are preparing for work with an international community can examine the available data and experience a taste of the community based even on the type of data that is collected. A well-intentioned and invested decision-maker will take the necessary steps to create a successful project atmosphere and outcome, they just need the guidance to do so effectively. It is the hope of this research that the value of the contextual process is evident, the nuances and challenges of relying on remote data are clear, and that the path to future research has been sufficiently paved.

5.1 Future Work

There a multitude of continuation opportunities to further explore the integration of big data and contextual engineering. This section highlights some opportunities that were considered during the development of this methodology but were ultimately outside of the scope of work.

As we operate in a world of global travel restrictions, the face-to-face limitations are known and familiar to nearly all parties across the world. This poses the thought that this mutual understanding has the potential to benefit the remote application of a traditionally in-person process such as contextual engineering. There is hope that future work can be completed with this consideration in mind by examining the willingness for stakeholders to meet and conduct important project conversations virtually and the impact this process may have on potential project success. This proposal has potential to work hand-in-hand with the modified international engineering project design process suggested in Figure 1.1 as these virtual communications could bolster the recommended communications with the client through the project design and implementation cycles, as represented by the dashed orange box.

There additionally exists potential for further supplementary data to be incorporated into a remote data analysis, specifically remote qualitative data as there exists a plethora of literature and media both about rural communities and created by rural communities. For example, the village of Tikonko has a handful of journal articles and news publications written about it, some short stories and literature written by community members, and an active public Facebook group. The incorporation of additional data sources, especially qualitative ones, has the potential to fill some of the gaps observed by the strictly quantitative methodology.

References

Africa Information Highway. (2020a). *Data catalog - sierra leone data portal* [Open data for africa]. Retrieved October 30, 2020, from https://sierraleone.opendataforafrica.org/data/#region=SL-W

Africa Information Highway. (2020b). *Ghana data portal* [Open data for africa]. Retrieved March 28, 2021, from https://ghana.opendataforafrica.org/

African Development Bank. (2017, March). *Economic benefits of open data in africa.* African Development Bank. Retrieved October 28, 2020, from https://www.tralac.org/documents/resources/africa/1752-economic-benefits-of-open-data-in-africa-afdb-march-2017/file.html

African Development Bank. (2020). *Home - africa information highway portal* [Open data for africa]. Retrieved October 28, 2020, from https://dataportal.opendataforafrica.org/

Balodi, R., Vati, L., Bisht, S., & Ghatak, A. (2015). *Rural youth migration and measures to combat it efficiently through agriculture* (1st). Retrieved June 26, 2019, from https://www.researchgate.net/publication/278683390_Rural_Youth_Migration_and_Measures_to_Combat_it_Efficiently_through_Agriculture

Bouck, E. C. (2004). How size and setting impact education in rural schools [Number: 3]. *Rural Ed.*, *25*(3). https://doi.org/10.35608/ruraled.v25i3.528

Calefato, F., Lanubile, F., & Minervini, P. (2010). Can real-time machine translation overcome language barriers in distributed requirements engineering? [ISSN: 2329-6313]. *2010 5th IEEE International Conference on Global Software Engineering*, 257–264. https://doi.org/10.1109/ICGSE.2010.37

CDC. (2020, February 11). *Coronavirus disease 2019 (COVID-19)* [Centers for disease control and prevention]. Retrieved October 25, 2020, from https://www.cdc.gov/coronavirus/2019-ncov/travelers/travel-during-covid19.html

Cherry, S., Robinson, A., Jashinsky, J., Bagwell-Adams, G., Elliott, M., & Davis, M. (2017). Rural community health needs assessment findings: Access to care and mental health. *J. Soc. Behav. Health Sci.*, *11*, 268–277. https://doi.org/10.5590/JSBHS.2017.11.1.18

Chinazzi, M., Davis, J. T., Ajelli, M., Gioannini, C., Litvinova, M., Merler, S., Piontti, A. P. y., Mu, K., Rossi, L., Sun, K., Viboud, C., Xiong, X., Yu, H., Halloran, M. E., Longini, I. M., & Vespignani, A. (2020). The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak [Publisher: American Association for the Advancement of Science Section: Research Article]. *Science*, *368*(6489), 395–400. https://doi.org/10.1126/science.aba9757

Deotti, L., & Estruch, E. (2016). *Addressing rural youth migration at its root causes: A conceptual framework*. Food and Agriculture Organization of the United

Nations. Social Policies, Rural Institutions Division. Retrieved January 26, 2020, from http://www.fao.org/3/a-i5718e.pdf

Deutsche Welle. (2020). *Sierra leone's tech-aided war on corruption | DW | 17.12.2020* [DW.COM]. Retrieved March 4, 2021, from https://www.dw.com/en/sierra-leones-tech-aided-war-on-corruption/a-55897439

Duffy, J., Lord, S. M., & Tsang, E. (2000). Service learning in engineering: What, why, and how?, 9.

Engineers Against Poverty. (2021). *Our approach | engineers against poverty.* Retrieved April 2, 2021, from http://engineersagainstpoverty.org/our-approach/

EWB. (n.d.). *521-pre-assessment_report_instructions.pdf* [Engineers without borders USA]. Retrieved January 25, 2020, from http://www.ewbncsu.org/wiki/images/c/c2/521-Pre-Assessment_Report_Instructions.pdf

EWB. (2019, May 8). *ICP lessons learned* [Engineers without borders USA]. Retrieved April 18, 2020, from https://ewb-usa.force.com/VolunteerVillage/s/article/ICP-Lessons-Learned

EWB. (2020a). *ICP lesson learned explorer* [Engineers without borders USA].

EWB. (2020b). *Measuring impact* [Engineers without borders USA] [Library Catalog: www.ewb-usa.org]. Retrieved April 20, 2020, from https://www.ewb-usa.org/measuring-impact/

FAO. (2018). *The state of food and agriculture: Migration, agriculture and rural development.* Food and Agriculture Organization of the United Nations.

Retrieved January 25, 2020, from http://www.fao.org/3/I9549EN/i9549en.pdf

IPUMS. (2019, February 7). *What is IPUMS?* [IPUMS]. Retrieved October 27, 2020, from https://ipums.org/what-is-ipums

Jordan, W. (2015). Creating a sustainable humanitarian engineering program. *2015 IEEE Canada International Humanitarian Technology Conference (IHTC2015)*, 1–4. https://doi.org/10.1109/IHTC.2015.7238040

Knoema. (2020). *Free data, statistics, analysis, visualization & sharing - knoema.com* [Knoema]. Retrieved October 28, 2020, from https://knoema.com//

Lee, T., Berg, D. R., & Buchanan, E. A. (2018). Challenges and opportunities in international service learning. *Proceedings of the ASEE Annual Conference & Exposition*, 1–15. Retrieved January 26, 2020, from http://search.ebscohost.com/login.aspx?direct=true&db=asn&AN=131758626

Litchfield, K., & Javernick-Will, A. (2014). Investigating gains from EWB-USA involvement. *J. Prof. Issues Eng. Ed. Pr.*, *140*(1), 04013008. https://doi.org/10.1061/(ASCE)EI.1943-5541.0000181

Lohr, S. (2012). The age of big data. *The New York Times*, 5.

Madden, S. (2012). From databases to big data [Conference Name: IEEE Internet Computing]. *IEEE Internet Comput.*, *16*(3), 4–6. https://doi.org/10.1109/MIC.2012.50

Martindale, T. (2013). Planning, monitoring, evaluation and learning program description. *Engineers Without Borders, USA*, 51.

Nations Online. (2014). *Political map of sierra leone - nations online project* [Nations online]. Retrieved October 26, 2020, from https://www.nationsonline.org/oneworld/map/sierra_leone_map2.htm

OpenStreetMap contributors. (2021). *OpenStreetMap* [OpenStreetMap]. Retrieved March 25, 2021, from https://www.openstreetmap.org/

Parker, K., Horowitz, J. M., Brown, A., Fry, R., Cohn, D., & Igielnik, R. (2018, May 22). *How urban, suburban and rural communities view one another* [Pew research center's social & demographic trends project]. Retrieved April 2, 2021, from https://www.pewresearch.org/social-trends/2018/05/22/how-people-in-urban-suburban-and-rural-communities-see-each-other-and-say-others-see-them/

Pateman, T. (2011). Rural and urban areas: Comparing lives using rural/urban classifications. *Region. Trends*, *43*(1), 11–86. https://doi.org/10.1057/rt.2011.2

Pew Research Center. (2016). *Religions in sierra leone | PEW-GRF* [Global religious futures]. Retrieved August 6, 2020, from http://www.globalreligiousfutures.org/countries/sierra-leone#/?affiliations_religion_id=0&affiliations_year=2010&region_name=All%20Countries&restrictions_year=2015

Reed, T., & Robinson, J. A. (2013). *The chiefdoms of sierra leone.*

Ridgway, C. (2015). Learning lessons at EWB-USA, 43.

Ritchie, H. (2018, September 27). *How urban is the world?* [Our world in data]. Retrieved February 26, 2020, from https://ourworldindata.org/how-urban-is-the-world

Schreiber, K. L., Rodríguez, L. F., Witmer, A.-P., & Dill, B. (2019). Understanding and incorporating stakeholder perspectives in international engineering: A phrase mining analysis. *2019 Boston, Massachusetts July 7- July 10, 2019*. https://doi.org/10.13031/aim.201901425

Statistics Sierra Leone. (2015). *Nuptiality and fertility in sierra leone - sierra leone data portal* [Open data for africa]. Retrieved March 28, 2021, from https://sierraleone.opendataforafrica.org/wiceqyg/nuptiality-and-fertility-in-sierra-leone

Statistics Sierra Leone. (2016). *2015 population and housing census: Summary of final results*. Freetown, Sierra Leone. Retrieved April 15, 2020, from https://www.statistics.sl/images/StatisticsSL/Documents/final-results_-2015_population_and_housing_census.pdf

Statistics Sierra Leone. (2020). *Stats SL - home | statistics sierra leone*. Retrieved October 28, 2020, from https://www.statistics.sl/

United Nations. (2017). *United nations statistics division - demographic and social statistics* [UN stats]. Retrieved April 2, 2021, from https://unstats.un.org/unsd/demographic/sconcerns/densurb/densurbmethods.htm

United Nations. (2018a, May 16). *68% of the world population projected to live in urban areas by 2050, says UN* [UN DESA | united nations department of economic and social affairs]. Retrieved January 25, 2020, from https://www.un.org/development/desa/en/news/population/2018-revision-of-world-urbanization-prospects.html

United Nations. (2018b). *World urbanization prospects: The 2018 revision*. United Nations Department of Economic and Social Affairs Population Division. Retrieved April 15, 2020, from https://population.un.org/wup/

United States Census. (2020). *Census.gov*. Retrieved October 28, 2020, from https://www.census.gov/

United States Department of State. (2019). SIERRA LEONE 2019 INTERNATIONAL RELIGIOUS FREEDOM REPORT. *Human Rights*, 5.

VSL Travel. (2021). *Explore bonthe and sherbro island* [Visit sierra leone (VSL TRAVEL)]. Retrieved March 25, 2021, from https://www.visitsierraleone.org/tour/explore-bonthe-sherbro-island/

Walters, J. P., & Javernick-Will, A. N. (2015). Long-term functionality of rural water services in developing countries: A system dynamics approach to understanding the dynamic interaction of factors. *Envir. Sci. Tech.*, *49*(8), 5035–5043. https://doi.org/10.1021/es505975h

Wang, Z., Li, D., Cheng, & Luo, T. (2020). Multifaceted influences of urbanization on sense of place in the rural–urban fringes of china: Growing, dissolving, and transitioning. *Journal of Urban Planning and Development*, *146*(1), 04019026. https://doi.org/10.1061/(ASCE)UP.1943-5444.0000540

Witmer, A.-P. (2018a). Addressing the influence of context and development in rural international engineering design.

Witmer, A.-P. (2018b). Contextual engineering assessment using an influence-identification tool. *J. Eng. Design Tech.*, *16*(6), 889–909. https://doi.org/10.1108/JEDT-05-2018-0091

World Bank. (n.d.). *About us | data*. Retrieved October 26, 2020, from https://data.worldbank.org/about

World Bank. (2019). *Mortality rate, under-5 (per 1,000 live births) - sierra leone | data*. Retrieved March 29, 2021, from https://data.worldbank.org/indicator/SH.DYN.MORT?locations=SL

World Bank. (2020). *Sierra leone | data*. Retrieved October 28, 2020, from https://data.worldbank.org/country/sierra-leone

APPENDIX A

INSTITUTIONAL REVIEW BOARD LETTER

ILLINOIS

OFFICE OF THE VICE CHANCELLOR FOR RESEARCH

Office for the Protection of Research Subjects
805 W. Pennsylvania Ave., MC-095
Urbana, IL 61801-4822

Notice of Approval: Amendment 03

March 11, 2019

Principal Investigator	Ann-Perry Witmer
Protocol Title	*International Engineering Effectiveness*
Protocol Number	17695
Funding Source	Unfunded
Review Type	Expedited 6, 7
Amendment Requested	• Reinstatement of subjects previously removed by amendment • Updating study locations: Removing Llano Largo, Honduras Adding El Durazno, Guatemala (study dates March 16 – 25, 2019) & Kenema, Sierra Leone (study dates May 30 – June 8, 2019) • Updating research team: Removing Kelsey Schreiber, Grace Witmer, Amber Kainz, & Michael Stablein Adding Julissa Nunez, Alexandra Timmons, Nicholas Perozzi, & Keilin Jahnke
Status	Active
Risk Determination	No more than minimal risk
Approval Date	March 11, 2019
Expiration Date	April 26, 2020

This letter authorizes the use of human subjects in the above protocol. The University of Illinois at Urbana-Champaign Institutional Review Board (IRB) has reviewed and approved the research study as described.

The Principal Investigator of this study is responsible for:

- Conducting research in a manner consistent with the requirements of the University and federal regulations found at 45 CFR 46.
- Using the approved consent documents, with the footer, from this approved package.
- Requesting approval from the IRB prior to implementing modifications.
- Notifying OPRS of any problems involving human subjects, including unanticipated events, participant complaints, or protocol deviations.
- Notifying OPRS of the completion of the study.

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Figure A.1: IRB authorization for use of human subjects in the detailed protocol.

APPENDIX B

REMOTE AND ON-SITE INDICATORS

Variable	Indicator Name	Source
CU2	CPIA policies for social inclusion/equity cluster average (1=low to 6=high)	World Bank Group, CPIA database (http://www.worldbank.org/ida).
CU3	CPIA gender equality rating (1=low to 6=high)	World Bank Group, CPIA database (http://www.worldbank.org/ida).
CU4	Bribery incidence (% of firms experiencing at least one bribe payment request)	World Bank, Enterprise Surveys (http://www.enterprisesurveys. org/).
CU5	CPIA social protection rating (1=low to 6=high)	World Bank Group, CPIA database (http://www.worldbank.org/ida).
CU6	Agricultural land (% of land area)	Food and Agriculture Organization, electronic files and web site.
EC1	Adjusted net national income per capita (current US$)	World Bank staff estimates based on sources and methods in World Bank's "The Changing Wealth of Nations: Measuring Sustainable Development in the New Millennium" (2011).
EC2	Maternal mortality ratio (modeled estimate, per 100,000 live births)	WHO, UNICEF, UNFPA, World Bank Group, and the United Nations Population Division. Trends in Maternal Mortality: 1990 to 2015. Geneva, World Health Organization, 2015
ED2	Literacy rate, adult female (% of females ages 15 and above)	United Nations Educational, Scientific, and Cultural Organization (UNESCO) Institute for Statistics.
ED1	Literacy rate, adult male (% of males ages 15 and above)	United Nations Educational, Scientific, and Cultural Organization (UNESCO) Institute for Statistics.
ED4	Primary completion rate, female (% of relevant age group)	United Nations Educational, Scientific, and Cultural Organization (UNESCO) Institute for Statistics.
ED3	Primary completion rate, male (% of relevant age group)	United Nations Educational, Scientific, and Cultural Organization (UNESCO) Institute for Statistics.
ED6	School enrollment, tertiary (% gross)	United Nations Educational, Scientific, and Cultural Organization (UNESCO) Institute for Statistics.
ED5	School enrollment, secondary (% gross)	United Nations Educational, Scientific, and Cultural Organization (UNESCO) Institute for Statistics.

Figure B.1: Expanded definitions of the 25 indicators used in Witmer's country influence calibration calculations, their associated variable names, and their source (Witmer, 2018b). These indicators were additionally purposed in determining the indicators to be utilized for the case study evaluation.

Variable	Indicator Name	Source
PO1	CPIA property rights and rule-based governance rating (1=low to 6=high)	World Bank Group, CPIA database (http://www.worldbank.org/ida).
PO2	CPIA public sector management and institutions cluster average (1=low to 6=high)	World Bank Group, CPIA database (http://www.worldbank.org/ida).
TE1	Access to electricity (% of population)	World Bank, Sustainable Energy for All (SE4ALL) database from the SE4ALL Global Tracking Framework led jointly by the World Bank, International Energy Agency, and the Energy Sector Management Assistance Program.
PO3	CPIA quality of public administration rating (1=low to 6=high)	World Bank Group, CPIA database (http://www.worldbank.org/ida).
PO4	CPIA transparency, accountability, and corruption in the public sector rating (1=low to 6=high)	World Bank Group, CPIA database (http://www.worldbank.org/ida).
TE3	Fixed broadband subscriptions (per 100 people)	International Telecommunication Union, World Telecommunication/ICT Development Report and database.
TE4	Renewable electricity output (% of total electricity output)	World Bank, Sustainable Energy for All (SE4ALL) database from the SE4ALL Global Tracking Framework led jointly by the World Bank, International Energy Agency, and the Energy Sector Management Assistance Program.
EC3	Inflation, consumer prices (annual %)	International Monetary Fund, International Financial Statistics and data files.
PO5	Government Effectiveness: Estimate	Detailed documentation of the WGI, interactive tools for exploring the data, and full access to the underlying source data available at www.govindicators.org.The WGI are produced by Daniel Kaufmann (Natural Resource Governance Institute and Brookings Institution) and Aart Kraay (World Bank Development Research Group). Please cite Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi (2010). "The Worldwide Governance Indicators: Methodology and Analytical Issues". World Bank Policy Research Working Paper No. 5430
PO6	Political Stability and Absence of Violence/Terrorism: Estimate	Detailed documentation of the WGI, interactive tools for exploring the data, and full access to the underlying source data available at www.govindicators.org.The WGI are produced by Daniel Kaufmann (Natural Resource Governance Institute and Brookings Institution) and Aart Kraay (World Bank Development Research Group). Please cite Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi (2010). "The Worldwide Governance Indicators: Methodology and Analytical Issues". World Bank Policy Research Working Paper No. 5430
TE2	Improved sanitation facilities (% of population with access)	WHO/UNICEF Joint Monitoring Programme (JMP) for Water Supply and Sanitation (http://www.wssinfo.org/).
CU1	Religious Composition by Country 2010-2050	Pew Research Center (http://www.pewforum. org/2015/04/02/religious-projection-table/#)

Figure B.1: (cont.)

Checklist of Contextual Influences

Location:
Date:
Assessor:

Guidelines
1 = no, not present, never occurs
3 = sometimes, present but not prevalent, occurs occasionally
5 = yes, prevalent, occurs often

	SCORE
Have you observed at least 25% of the community's residents visibly manifesting a cultural identification beyond the national standard through language, dress, behavior, or religious practice?	
Do community members who manifest a cultural identification live and work fully integrated with segments of community population that do not share cultural identification? (5 if no cultural diversity)	
Do residents practice behaviors, social habits, politics, etc. that they overtly attribute to spiritual beliefs?	
Do residents use technology or technical practices that they say are unique to their cultural identification or history?	
Is there a primary religion practiced in the community?	
Do community members overtly express discrimination against anyone based on physical characteristics, cultural practices, or religious beliefs?	
Do the majority of younger adults follow dress, language and/or behaviors associated with the cultural traditions of their elders?	
Does the community elect its leadership by participation of all residents?	
Can residents accurately describe how their local government is organized and managed?	
Do attendees in community public meetings disagree with leaders and express personal opinions, even if they differ from the majority opinion?	
Is there an effective and widely known process whereby residents may effectively air grievances and/or seek redress from their community's leadership?	
Does the community have sub-groups of residents, such as neighborhoods, family pods, or coalitions?	
Does any resident have the equal right to participate in a governing board?	
Can decisions made by local governance be overruled by a higher level of government?	
Does the community use written documentation in governance?	
Have you witnessed any violent disagreements or threat of violence, either during community meetings or in public conversations outside of meetings?	
What is the average apparent age of community leaders? >90% NLE=1, >80% NLE=2, >70% NLE=3, >60%=4, >50%=5	
Is there evidence of infrastructure projects that have been provided in past or present by international organizations?	
Do community leaders appear to benefit from leadership through money/bribes, status, social control, or living comfort?	
Is the community's population self-employed? >80%=5, >60%=4,>40%=3,>=20%=2, <20%=1	
Do farmers in the community sell their products to others? >70%=1, >60%=2, >50%=3, >40%=4, <=40%=5	
Are there stores, banks, and other service outlets located within community boundaries?	
Are homes built using materials and methods typically seen in larger cities in the country? >70%=5, >55%=4, >40%=3, >25%=2, <=10%=1	
Are streets in the community's population center paved? >90%=5, >70% = 4, >50%=3, >30%=2, <=30%=1	
Is there a health clinic or hospital located within the community or immediately adjacent and easily accessible?	
What is the apparent average age of the community? >=70% NLE=1, <70% NLE=2, <60% NLE=3, <50% NLE=4, <40% NLE=5	
Do you see upholstered furniture in homes in the community? Always=5, Usually=4, Sometimes=3, Rarely=2, Never=1	
Are there sanitary latrines in or easily accessible to homes? Always=5, Usually=4, Sometimes=3, Rarely=2, Never=1	
Do you see heavy alcohol use outside of social or ceremonial gatherings (ie alcohol consumption at workplace, public drunkenness)? Always=5, Often=4, Sometimes=3, Rarely=2, Never=1	
Do homes in the community have elecricity? Always=5, Usually=4, Sometimes=3, Rarely=2, Never=1	
Do most residents own a cell phone? Always=5, Usually=4, Sometimes=3, Rarely=2, Never=1	
Is there refrigeration in residences? Always=5, Usually=4, Sometimes=3, Rarely=2, Never=1	
Do residents demonstrate an interest in understanding or obtaining visiting technology they're not familiar with? Always=5, Usually=4, Sometimes=3, Rarely=2, Never=1	
Do residents own motorized vehicles? Always=5, Usually=4, Sometimes=3, Rarely=2, Never=1	
Do residents correlate illness with water-borne disease? Always=5, Usually=4, Sometimes=3, Rarely=2, Never=1	
Is there an elementary school in or near town that children from the community may easily attend?	
Is there a secondary school in or near town that children from the community may easily attend?	
Do community children attend primary school for general education? >90%=5, >70% = 4, >50%=3, >30%=2, <=30%=1	
Do community children attend secondary school for general education? >90%=5, >70% = 4, >50%=3, >30%=2, <=30%=1	
Do children from the community attend university? >40%=5, >30%=4, >20%=3, >10%=2, <=10%=1	
Are community leaders educated? Often through university=5, Often through secondary =4, sometimes through secondary=3, often through primary=2, sometimes through primary=1	

Figure B.2: Predictive Tool created by Witmer for on-site contextual data collection by technical designers working in international communities (Witmer, 2018b).

APPENDIX C

INDICATOR MATCHING

For further information on the indicator matching process, please visit the Indicator Matching spreadsheet available here: https://bit.ly/2PsRYn5.